

Looking Worse Before We Look Better Part II: Communicating 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

It is widely accepted that the adoption and implementation of the Common Core State Standards will raise the academic bar of rigor and demand for U.S. students. Most have reasonably assumed that the accompanying common core assessments—Partnership for Assessment of Readiness for College and Careers (PARCC) and Smarter Balanced Assessment Consortium (SBAC)—will reflect that increased rigor. While there is a healthy sense of optimism building around the release of these two assessments, there is a general concern that what constitutes “proficiency” on these two assessments will be far more rigorous than what most states have previously set. One likely, unintended consequence of the release of the new assessments is that states around the country will appear to have suffered a precipitous decline in academic progress (Smith & Turner, 2013). Divorced from context, however, those assessment scores threaten to paint an inaccurate picture of student growth.

As an example of how assessment results may be misconstrued, consider a state that reports 75 percent of their third graders were proficient in reading in 2013. Then, in 2014, the state reports that (under the new assessment consortium program) only 50 percent are proficient. To the uninitiated, it appears that the state suffered a marked decline in reading growth, when in fact, it is the benchmark for what constitutes ‘proficiency,’ along with introduction of a common scale, that has been raised and is responsible for the perceived drop in student performance.

Fear that the transition to a new set of assessments may initially present an inaccurate picture of state educational progress, a picture that makes for easy headline grabbing, is nothing new. In fact, Wayne Camara and Emily Shaw made a compelling case that when it comes to reporting on educational test data, mainstream media sources, more often than not, cite the opinions of “testing opponents as the sole commentary in their articles on testing policy” (Camara & Shaw, 2012). In terms of the general public, a 2010 survey found that 75% receive their news on education topics from families and friends (Camara &

Shaw, 2012). Meaning, perception matters. The perception that a state’s academic performance is suddenly sliding in the wrong direction can be devastating to educators at all levels, from the classroom to the district office to the state department of education. Educational policy makers are well aware of this public perception issue. As a recent article, *The Schoolmaster*, in *The Atlantic* makes clear, even one of the primary authors of the Common Core, David Coleman, admits the transition to PARCC and SBAC will initially be perceived as an academic stumble:

Coleman admits that the Core will probably lead to “a short-term reduction in [test] scores,” but he seems to have made peace with this reality as a necessary hardship on the road to his academic utopia.

In that same article, Jeb Bush, a well-known advocate for education reform, admits the terrible political pressure that state and educational leaders will face in the coming years:

“The big fight will be coming in 2014, when we begin to implement and assess these standards. If a third are ready, what will the response of states be then? Will they do what they historically have done, which is to pull back and say ‘Oh my God, that’s not fair, excuse, excuse, excuse?’ Or will they accept responsibility to say ‘That’s the fact, that’s where we are now. Maybe 40 percent of our kids are ready if we benchmark them against the world?’”

Bush’s tough-love position is easy for a former politician to take, but less so for a current elected official. After all, no governor or legislator wants to preside over plummeting test scores. Pressure to roll back the Common Core or to relax the tests may be intense.

This perception, of course, is unfair. Most educators expect student assessment measures to decline under the new assessment scheme. This is not necessarily because student *performance* is actually

For more information on Lexile measures and the Common Core State Standards, visit www.Lexile.com.

falling or because students are actually performing worse. It is merely because the educational bar used to measure student performance, the benchmark for proficiency, has risen. However, these more sophisticated and nuanced explanations make for lousy sound bites. For the uninitiated public (including mainstream media), one unfamiliar with basic principles of assessment and psychometrics, a much more attention-grabbing story is that the educational sky is falling. Thus, educators nationwide are bracing for the impending backlash.

Of course, it does not have to be that way. Stephen Sireci has argued that even abstract philosophical and educational issues can be communicated to mainstream media in more palatable ways. Sireci recommends employing three strategies; the use of plain language to convey assessment information, the use of visuals, and the use of stories and examples (Sireci & Forte, 2012). Camara and Shaw demonstrate a similar understanding in their call for education departments to engage with the mainstream media:

It would also be worthwhile to consider holding a pre-session with the media if you are introducing a new assessment or if there will be major changes to the test. If the media are not familiar with the assessment or there have been major changes in the assessment it is likely they will have preconceived ideas about the results, which are easier to address before results are released and reporters are on a deadline.

The challenge for educators is to proactively manage public perceptions *before* the release of the assessment results, and to communicate how the transition to the common core creates a higher bar for U.S. students. 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. Similarly, 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 adoption of our new 'national assessments' and what that adoption will mean for the measurement of growth.

Following Sireci's recommendations, part of that state effort to effectively communicate should involve the use of clear analogies and examples. Analogies are rhetorically powerful and, when used properly, can be instructive and illuminating. Consider a simple analogy taken from track and field. Imagine a track coach who sets a proficiency bench-

mark of three feet for her hurdle jumpers. Using a simple process, she is able to identify the percentage of athletes who are able to clear three feet and are thus labeled proficient. Now imagine a new year brings a new coach who feels that three feet is no longer a competitive height in the world of track and field and so updates the proficiency benchmark to three feet, six inches. Because the bar has been raised (literally in this case), a fewer number of athletes may now qualify as proficient. However, to an outside observer it appears that the percentage of proficient track and field athletes has declined. In fact, the athletes have the same ability they had before, and it is unlikely that any athlete actually lost athletic ability. It is simply that what we take to constitute proficiency has been altered.

The common core assessments present educators with the same bitter but necessary pill: if we want U.S. students to graduate college and career ready in an increasingly global and competitive world, then more rigorous academic standards are required. That rigor demands that we raise the proficiency benchmark for all our students in the interest of long term gains. In the short term, however, transitioning to the new common core assessments may give rise to the perception that our educational progress has faltered.

It could have been different. Such a communication nightmare could have been avoided if over the life of *No Child Left Behind* (NCLB) we had taken a growth perspective and not a status perspective in describing student achievement. State departments of education could have focused on longitudinal growth over cross-sectional growth, and used vertical scales. 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.

To punctuate the problem with a temperature analogy, imagine a world in which weather reporters did not report the actual temperature but just categorical labels (hot, warm, cool, cold). What Floridians consider "cold" would no doubt be very different from the actual temperature that a Minnesotan would report as being cold. However, even in this hypothetical world of temperature-free weather reporting, there is an obvious path away from the confusion:

underlying these descriptive labels are vertical and exchangeable scales (Fahrenheit and Celsius). That is, we could explain that Floridians feel like it is cold when the temperature is 48°F and below, whereas Minnesotans consider it cold when it is 32°F and below. Even if weather forecasters in Minnesota used the Celsius scale instead of Fahrenheit, because the two scales are exchangeable, we could still impose order on these different perceptions.

By allowing each state to define its own level of proficiency, we created an academic Tower of Babel of reading achievement in this country. Unfortunately, the basic measurement principles that are commonplace and essential in the “hard sciences” have been for too long ignored in the “soft sciences.” However, when it comes to the measurement of reading, that neglect is no longer justified. The Lexile® Framework for Reading is a vertical, developmental scale that allows the measurement of students and text on a common scale.

Today, there are 20 states that report Lexile® measures on their NCLB assessment. For the states already employing the Lexile metric, the message can be clear and concise as they transition to the more rigorous demands of the CCSS. To revisit the hurdle analogy, a state can report that last year’s proficiency hurdle was a Lexile level of X, but that hurdle has been raised to Y. Additionally, the state can report that while the hurdle has been raised, student performance did not actually slide, and that the average Lexile measure for students is (hopefully) higher than the previous year. In other words, state officials can explain that their students are jumping as high (or higher) than last year, but we have simply raised the hurdle requirement. These sorts of clear analogies simplify the message and allow for better understanding.

Additionally, the fact that the Lexile scale can also be used in the measurement of text complexity provides students, parents and educators with a concrete, interpretive framework. This feature of the conjoint measurement of reading comprehension and text complexity is a critical component in realizing the full potential of the CCSS. As Standard 10 in the Common Core State Standards for English Language Arts/Literacy Appendix A indicates, the ability of students to read increasingly complex texts is one of the major shifts from previous standards (National Governor’s Association, 2012). As ubiquitous as Lexile measures are in the measurement of student reading ability, on the text side, Lexile measures are even more pervasive. Today, every major text and trade book publisher have measured their books using the Lexile Frame-

work. In addition, millions of articles have also been measured and are freely available from library resource groups like EBSCO and ProQuest.

The total and complete implementation of CCSS will require incredible commitment, time and persistence. To facilitate this Herculean effort, our organization has provided frictionless and free resources to educators as they align and build new resources. This effort has resulted in over 100,000 educators, parents and students measuring over 700 million words through the fourth quarter of 2012.

The conjoint feature of measuring text and students on the Lexile scale, combined with the fact that the CCSS identified the Lexile level per grade for college and career readiness, provides students, parents and educators with clear empirical and external benchmarks. No longer will students, parents and educators have to guess and wait for some annual assessment to determine if they are on track for college and career readiness. Furthermore, like the adoption of the CCSS, the adoption and acceptance of Lexile measures has been a voluntary movement spawned by educators that recognized the utility and value of the measures for their classroom practice.

It may seem unimaginable to us now, but the unification of measurement has been a slow road. There was a time in our history when we did not have unification of the measurement of constructs such as temperature and time. In the early 1600s there were dozens of instrument makers measuring temperature on their own scale. However, once our scientific understanding of temperature was established, there was recognition that we did not need all these different instrument makers, but common scales. Additionally, as late as the 1850s, we did not have standardized time. A Canadian engineer, Sanford Fleming, frustrated with translating all the different railroad time tables in order to figure out when his nephew would arrive at the train station, convened a world summit to fix this problem by creating standard time zones.

Now the time has come in education that we learn the lessons from the measurement of temperature and time, and realize the importance of the unification of measurement. The education of our students will be best facilitated not by adding more tests with various scales, but by placing the tests that we administer on a common scale. Because the Lexile Framework for Reading is a vertical scale used universally by publishers and educators, the benefits of measuring longitudinal growth across an array of assessments is now possible. The conjoint feature of the measurement of text and individuals on the same scale permits a rich,

interpretive framework for understanding test scores. The text demands for college and career readiness expressed in Lexile measures gives students, parents and educators a clear, transparent and empirical benchmark upon which to evaluate their progress. The late Steven Covey stated that one of the core principles of effective leaders is to “begin with the end in mind.” This core tenet, when it comes to college and career readiness for reading, has now been identified through the CCSS and expressed by grade level. As we transition to these higher and more rigorous demands, the Lexile Framework provides educators and policy makers with another tool in their arsenal, not only for achieving these goals, but for communicating progress. Along with the analogy of the higher hurdle of the CCSS, we can reassure the public that student achievement in reading is not getting worse, but that what constitutes our new standards has become more rigorous. Through the Lexile scale for reading we can show that student performance is growing even as we transition to the more rigorous standards. Ultimately the success of the CCSS may rest on how well we communicate student test performance in this

transitional moment, which at first glance, may look worse than it actually is.

REFERENCES

Camara, W. J., & Shaw, E. J. (2012). The Media and Educational Testing: In Pursuit of the Truth or in Pursuit of a Good Story? *Educational Measurement: Issues and Practice*, 31, 33-37.

Goldstein, D. (2012). The Schoolmaster. *The Atlantic*.

National Governor’s Association, Council of Chief State School Officers. (2012). *Common Core State Standards Initiative: Standard 10: Range, Quality, & Complexity*. Retrieved from Common Core State Standards Initiative: <http://www.corestandards.org/ELA-Literacy>

Sireci, S. G., & Forte, E. (2012). Informing in the Information Age: How to Communicate Measurement Concepts to Education Policy Makers. *Educational Measurement: Issues and Practice*, 31, 27-32.

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

Smith, M., Turner, J. (2013). Looking Worse Before We Look Better: Student Performance in a Common Core World. Durham, North Carolina: MetaMetrics, Inc.

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.

