# A letter to the U.S. Department of Education (updated July 14)

July 12, 2016


The Honorable John King

Secretary of the Education Department

400 Maryland Avenue, SW

Washington, D.C. 20202


Dear Mr. Secretary:

The Every Student Succeeds Act (ESSA) marks a great opportunity for states to advance accountability systems beyond those from the No Child Left Behind (NCLB) era. The Act (Section 1111(c)(4)(B)(i)(I)) requires states to use an indicator of academic achievement that "measures proficiency on the statewide assessments in reading/language arts and mathematics." The proposed rulemaking (§ 200.14) would clarify this statutory provision to say that the academic achievement indicator must "equally measure grade-level proficiency on the reading/language arts and mathematics assessments."

**We write this letter to argue that the Department of Education should not mandate the use of proficiency rates as a metric of school performance under ESSA.** That is, states should not be limited to measuring academic achievement using performance metrics that focus only on the proportion of students who are grade-level proficient—rather, they should be encouraged, or at a minimum allowed, to use performance metrics that account for student achievement at all levels, provided the state defines what performance level represents grade level proficiency on its reading/language arts and mathematics assessments.

Moving beyond proficiency rates as the sole or primary measure of school performance has many advantages. For example, a narrow focus on proficiency rates incentivizes schools to focus on those students near the proficiency cut score, while an approach that takes into account all levels of performance incentivizes a focus on *all* students. Furthermore, measuring performance using the full range of achievement provides additional and useful information for parents, practitioners, researchers, and policymakers for the purposes of decisionmaking and accountability, including more accurate information about the differences among schools.

Reporting performance in terms of the percentage above proficient is problematic in several important ways. Percent proficient:

1. Incentivizes schools to focus only on students around the proficiency cutoff rather than all students in a school (Booher-Jennings, 2005; Neal & Schanzenbach, 2010). This can divert resources from students who are at lower or higher points in the achievement distribution, some of whom may need as much or more support than students just around the proficiency cut score (Schwartz, Hamilton, Stecher, & Steele, 2011). This has been shown to influence which students in a state benefit (i.e., experience gains in their academic achievement) from accountability regulations (Neal & Schanzenbach, 2010).
2. Encourages teachers to focus on bringing students to a minimum level of proficiency rather than continuing to advance student learning to higher levels of performance beyond proficiency.
3. Is not a reliable measure of school performance. For example, percent proficient is an inappropriate measure of progress over time because changes in proficiency rates are unstable and measured with error (Ho, 2008; Linn, 2003; Kane & Staiger, 2002). The percent proficient is also dependent upon the state-determined cut score for proficiency on annual assessments (Ho, 2008), which varies from state to state and over time. Percent proficient further depends on details of the testing program that shouldn't matter, such as the composition of the items on the state test or the type of method used to set performance standards. These problems are compounded in small schools or in subgroups that are small in size.
4. Is a very poor measure of performance gaps between subgroups, because percent proficient will be affected by how a proficiency cut score on the state assessments is chosen (Ho, 2008; Holland, 2002). Indeed, prior research suggests that using percent proficient can even *reverse the sign* of changes in achievement gaps over time relative to if a more accurate method is used (Linn, 2007).
5. Penalizes schools that serve larger proportions of low-achieving students (Kober & Riddle, 2012) as schools are not given credit for improvements in performance other than the move to proficiency from not-proficient.

We suggest two practices for measuring achievement that lessen or avoid these problems. Importantly, some of these practices were utilized by states in ESEA Flexibility Waivers and are improvements to NCLB practices (Polikoff, McEachin, Wrabel, & Duque, 2014).

**Average Scale Scores**

The best approach for measuring student achievement levels for accountability purposes under ESSA is to use average scale scores. Rather than presenting performance as the proportion of students who have met the minimum-proficiency cut score, states could present the average (mean) score of students within the school and the average performance of each subgroup of students. If the Department believes percent proficient is also important for reporting purposes, these values could be reported alongside the average scale scores.

The use of mean scores places the focus on improving the academic achievement of *all* students within a school and not just those whose performance is around the state proficiency cut score

(Center for Education Policy, 2011). Such a practice also increases the amount of variation in school performance measures each year, providing for improved differentiation between schools that may have otherwise similar proficiency rates. In fact Ho (2008) argues if a single rating is going to be used for reporting on performance, it should be a measure of the average performance because such measures incorporate the value of every score (student) into the calculation and the average can be used for more advanced analyses. The measurement of gaps between key demographic groups of students, a key goal of the ESSA law, is dramatically improved with the use of average scores rather than the proportion of proficient students (Holland, 2002; Linn, 2007).

**Proficiency Indexes**

If average scale scores cannot be used, a weaker alternative that is still superior to percent proficient would be to allow states to use proficiency indexes. Schools under this policy would be allocated points based on multiple levels of performance. For example, a state could identify four levels of performance on annual assessments: Well Below Proficient, Below Proficient, Proficient, and Advanced Proficient. Schools receive no credit for students Well Below Proficient, partial credit for students who are Below Proficient, full credit for students reaching Proficiency, and additional credit for students reaching Advanced Proficiency. Here we present an example using School A and School B.

Proficiency Index Example

| Proficiency Category | School A | | | School B | | |
|---|---|---|---|---|---|---|
| | (A) Points Per Student | (B) # of Students | (C) Index Points | (A) Points Per Student | (B) # of Students | (C) Index Points |
| Well Below Proficient | 0.0 | 27 | 0.0 | 0.0 | 18 | 0.0 |
| Below Proficient | 0.5 | 18 | 9.0 | 0.5 | 27 | 13.5 |
| Proficient | 1.0 | 33 | 33.0 | 1.0 | 26 | 26.0 |
| Advanced Proficient | 1.5 | 22 | 33.0 | 1.5 | 29 | 43.5 |
| Total | | 100 | 75.0 | | 100 | 83.0 |

NCLB Proficiency Rate: 55%     NCLB Proficiency Rate: 55%
ESSA Proficiency Index: 75     ESSA Proficiency Index: 83

Under NCLB proficiency rate regulations, both School A and School B would have received a 55% proficiency rate score. Using a Proficiency Index, the performance of these schools would no longer be identical. A state would be able to *compare* the two schools while simultaneously identifying *annual meaningful differentiation* in the performance of School A from that of School B. The hypothetical case presented here is not the only way a proficiency index can be

used. Massachusetts is one example of a state that has used a proficiency index for the purposes of identifying low-performing schools and gaps between subgroup of students (see: ESEA Flexibility Request: Massachusetts, page 32). These indexes are understandable for practitioners, family members, and administrators while also providing additional information regarding the performance of students who are not grade-level proficient.

The benefits of using such an index, relative to using the proportion of proficient students in a school, is that it incentivizes a focus on all students, not just those around an assessment's proficiency cut score (Linn, Baker, & Betebenner, 2002). Moreover, schools with large proportions of students well-below the proficiency cut score are given credit for moving students to higher levels of performance even if still below the cut score (Linn, 2003). The use of a proficiency index or providing schools credit for students at different points in the achievement distribution improves the construct validity of the accountability measures over the NCLB proficiency rate measures (Polikoff et al., 2014). In other words, the inferences made about schools (e.g., low-performing or bottom 5%) using the proposed measures are more appropriate than those made using proficiency rates alone.

**What We Recommend**

Given the findings cited above, we believe the Department of Education should revise its regulations to one of two positions:

- Explicitly endorsing or encouraging states to use one of the two above-mentioned approaches as an alternative to proficiency rates as the primary measure of school performance. Average scale scores is the superior method.
- Failing that, clarifying that the law is neutral about the use of proficiency rates versus one of the two above-mentioned alternatives to proficiency rates as the primary measure of school performance.

With the preponderance of evidence showing that schools and teachers respond to incentives embedded in accountability systems, we believe option 1 is the best choice. This option leaves states the authority to determine school performance how they see fit but encourages them to incorporate what we have learned through research about the most accurate and appropriate way to measure school performance levels.

**Our Recommendation is Consistent with ESSA**

Section 1111(c)(4)(A)) of ESEA, as amended by ESSA, requires each state to establish long-term goals:

"(i) for all students and separately for each sub- group of students in the State—

(I) for, at a minimum, improved—

(aa) academic achievement, as measured by proficiency on the annual assessments required under subsection (b)(2)(B)(v)(I);"

And Section 1111(c)(4)(B) of ESEA requires the State accountability system to have indicators that are used to differentiate all public schools in the State, including—(i) "academic achievement—(I) as measured by proficiency on the annual assessments required [under other provisions of ESSA]."

Our suggested approach is supportable under these provisions based on the following analysis. The above-quoted provisions in the law that mandate long-term goals and indictors of student achievement based on proficiency on annual assessments do not prescribe how a state specifically uses the concept of proficient performance on the state assessments. The statute does not prescribe that "proficiency" be interpreted to compel differentiation of schools based exclusively on "proficiency rates." Proficiency is commonly taken to mean "knowledge" or "skill" (Merriam Webster defines it as "advancement in knowledge or skill" or "the quality or state of being proficient", where "proficient" is defined as "well advanced in an art, occupation, or branch of knowledge"). Under either of these definitions, an aggregate performance measure such as the two options described above would clearly qualify as involving a measure of proficiency. Both of the above-mentioned options provide *more* information about the average proficiency level of a school than an aggregate proficiency rate. Moreover, they address far more effectively than proficiency rates the core purposes of ESSA, including incentivizing more effective efforts to educate all children and providing broad discretion to states in designing their accountability systems.

We would be happy to provide more information on these recommendations at your pleasure.

Sincerely,

Morgan Polikoff, Ph.D., Associate Professor of Education, USC Rossier School of Education

Signatories

Educational Researchers

Alice Huguet, Ph.D., Postdoctoral Fellow, School of Education and Social Policy, Northwestern University

Andrew Ho, Ph.D., Professor of Education, Harvard Graduate School of Education

Andrew Saultz, Ph.D., Assistant Professor, Miami University (Ohio)

Andrew Schaper, Ph.D., Senior Associate, Basis Policy Research

Anna Egalite, Ph.D., Assistant Professor of Education, North Carolina State University

Arie van der Ploeg, Ph.D., retired Principal Researcher, American Institutes for Research

Cara Jackson, Ph.D., Assistant Director of Research & Evaluation, Urban Teachers

Christopher A. Candelaria, Ph.D., Assistant Professor of Public Policy and Education, Vanderbilt University

Cory Koedel, Ph.D., Associate Professor of Economics and Public Policy, University of Missouri

Dan Goldhaber, Ph. D., Director, Center for Education Data & Research, University of Washington Bothell

Danielle Dennis, Ph.D., Associate Professor of Literacy Studies, University of South Florida

Daniel Koretz, Ph.D., Henry Lee Shattuck Professor of Education, Harvard Graduate School of Education

David M. Rochman, Research and Program Analyst, Moose Analytics

Edward J. Fuller, Ph.D., Associate Professor of Education Policy, The Pennsylvania State University

Eric A. Houck, Associate Professor of Educational Leadership and Policy, University of North Carolina at Chapel Hill

Eric Parsons, Ph.D., Assistant Research Professor, University of Missouri

Erin O'Hara, former Assistant Commissioner for Data & Research, Tennessee Department of Education

Ethan Hutt, Ph.D., Assistant Professor of Education, University of Maryland College Park

Eva Baker, Ed.D., Distinguished Research Professor, UCLA Graduate School of Education and Information Studies, Director, Center for Research on Evaluation, Standards, and Student Testing, Past President, American Educational Research Association

Greg Palardy, Ph.D., Associate Professor, University of California, Riverside

Heather J. Hough, Ph.D., Executive Director, CORE-PACE Research Partnership

Jason A. Grissom, Ph.D., Associate Professor of Public Policy and Education, Vanderbilt University

Jeffrey W. Snyder, Ph.D., Assistant Professor, Cleveland State University

Joanne Weiss, former chief of staff to U.S. Secretary of Education Arne Duncan

John Q. Easton, Ph.D., Vice President, Programs, Spencer Foundation, former Director, Institute of Education Sciences

John Ritzler, Ph.D., Executive Director, Research & Evaluation Services, South Bend Community School Corporation

Jonathan Plucker, Ph.D., Julian C. Stanley Professor of Talent Development, Johns Hopkins University

Joshua Cowen, Ph.D., Associate Professor of Education Policy, Michigan State University

Katherine Glenn-Applegate, Ph.D., Assistant Professor of Education, Ohio Wesleyan University

Linda Darling-Hammond, Ed.D., President, Learning Policy Institute, Charles E. Ducommun Professor of Education Emeritus, Stanford University, Past President, American Educational Research Association

Lindsay Bell Weixler, Ph.D., Senior Research Fellow, Education Research Alliance for New Orleans

Martin R. West, Ph.D., Associate Professor, Harvard Graduate School of Education

Matt Chingos, Ph.D., Senior Fellow, Urban Institute

Matthew Di Carlo, Ph.D., Senior Fellow, Albert Shanker Institute

Matthew Duque, Ph.D., Data Strategist, Baltimore County Public Schools

Matthew A. Kraft, Ed.D., Assistant Professor of Education and Economics, Brown University

Michael Hansen, Ph.D., Senior Fellow and Director, Brown Center on Education Policy, Brookings Institution

Michael J. Petrilli, President, Thomas B. Fordham Institute

Nathan Trenholm, Director of Accountability and Research, Clark County (NV) School District

Russell Brown, Ph.D., Chief Accountability Officer, Baltimore County Public Schools

Sarah Reckhow, Ph.D., Assistant Professor of Political Science, Michigan State University

Sean P. "Jack" Buckley, Ph.D., Senior Vice President, Research, The College Board, former Commissioner of the National Center for Education Statistics

Sherman Dorn, Ph.D., Professor, Mary Lou Fulton Teachers College, Arizona State University

Stephani L. Wrabel, Ph.D., USC Rossier School of Education

Tom Loveless, Ph.D., Non-resident Senior Fellow, Brookings Institution

K-12 Educators

Alexander McNaughton, History Teacher, YES Prep Charter School, Houston, TX

Brett Coffman, Ed.S., Assistant Principal, Liberty High School, MO

Daniel Gohl, Chief Academic Officer, Broward County Public Schools

Danielle Blue, M.Ed., Director of Preschool Programming, South Kingstown Parks and Recreation, Rhode Island

Jacquline D. Price, M.Ed., County School Superintendent, La Paz County, AZ

Jillian Haring, Staff Assistant, Broward County Public Schools

Juan Gomez, Middle School Math Instructional Coach Carmel High School, Carmel, CA

MaryEllen Falvey, M.Ed, NBCT, Office of Academics, Broward County Public Schools

Oscar Garcia, Social Studies Teacher, El Paso Academy East

Patricia K. Hadley, Elementary School Teacher, Retired, Twin Falls, Idaho

Theodore A. Hadley, High School/Middle School Teacher, Retired, Twin Falls, Idaho
Troy Frystak, 5/6 Teacher, Springwater Environmental Sciences School


Other Interested Parties

Arnold F. Shober, Ph.D., Associate Professor of Government, Lawrence University

Celine Coggins, Ph. D., Founder and CEO, Teach Plus

Justin Reich, EdD, Executive Director, Teaching Systems Lab, MIT

Karl Rectanus, CEO, Lea(R)n, Inc.

Kenneth R. DeNisco, Ph.D., Associate Professor, Physics & Astronomy, Harrisburg Area Community College

Kimberly L. Glass, Ph.D., Pediatric Neuropsychologist, The Stixrud Group

Patrick Dunn, Ph.D., Biomedical Research Curator, Northrop Grumman TS

Steven Gorman, Ph.D., Program Manager, Academy for Lifelong Learning, LSC-Montgomery

Torrance Robinson, CEO, trovvit

References

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, *42*(1), 231–268.

Center on Education Policy. (2011, May 3). An open letter from the Center on Education Policy to the SMARTER Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Career. Retrieved from http://cep-dc.org/displayDocument.cfm?DocumentID=359

Ho, A. D. (2008). The problem with "proficiency": Limitations of statistics and policy under No Child Left Behind. *Educational Researcher*, *37*(6), 351–360.

Holland, P. W. (2002). Two measures of change in the gaps between the CDFs of test-score distributions. *Journal of Educational Behavioral Statistics*, *27*(1), 3–17.

Kober, N., & Riddle, W. (2012). *Accountability issues to watch under NCLB waivers*. Washington, DC: Center on Education Policy.

Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, *32*(7), 3–13.

Linn, R. L. (2007). *Educational accountability systems*. Paper presented at the The CRESST Conference: The Future of Test-Based Educational Accountability.

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, *31*(6), 3–16.

Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics*, *92*, 263–283.

Ng, H. L., & Koretz, D. (2015). Sensitivity of school-performance ratings to scaling decisions. *Applied Measurement in Education*, *28*(4), 330–349.

Polikoff, M. S., McEachin, A., Wrabel, S. L., & Duque, M. (2014). The waive of the future? School accountability in the waiver era. *Educational Researcher*, *43*(1), 45–54. http://doi.org/10.3102/0013189X13517137

Schwartz, H. L., Hamilton, L. S., Stecher, B. M., & Steele, J. L. (2011). *Expanded measures of school performance*. Santa Monica, CA: The RAND Corporation.