NEPC Review: High Stakes for High Achievers: State Accountability in the Age of ESSA and High Stakes for High Schoolers: State Accountability in the Age of ESSA

Beth C. Rubin
Rutgers University, beth.rubin@gse.rutgers.edu

Follow this and additional works at: https://scholar.colorado.edu/nepc
Part of the Education Commons

Recommended Citation

This NEPC Review is brought to you for free and open access by Centers and Research Institutes at CU Scholar. It has been accepted for inclusion in National Education Policy Center by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Two recent reports from the Fordham Institute address the question of the impact of state accountability systems on “high achievers,” referred to in the reports as “students who have already crossed the proficiency threshold.” Both reports argue that this group is being neglected educationally, and they advocate for accountability systems to be redesigned to attend to the needs of high-achieving students. Both reports also recommend that states use a “performance index,” as opposed to proficiency rates, to measure school achievement. This review, however, concludes that: 1) the reports’ central assumptions about high-achieving students are problematic; 2) growth measures are not an effective means for directing attention to high-achieving students; 3) narrow, high-stakes forms of assessment may negatively impact the education provided to these students; and, 4) further stratifying educational settings and reallocating resources toward “high-achieving” students has troublesome implications for the democratic goals of education. Implementation of the reports’ recommendations may in fact result in a furthering of the inequitable educational opportunities that ESSA was designed to reduce.
I. Introduction

The No Child Left Behind Act (NCLB) was signed into law in 2001, fueled by bipartisan concern about the standardized test scores of U.S. students in comparison with their international cohort and dissatisfaction with gaps in achievement between certain student groups – low income, English-language learners, special education, and students of color – and their peers. Passage of NCLB resulted in the development and application of standardized performance measures intended to both identify and incentivize school systems to address those gaps.

The law has been widely criticized for expanding federal influence over public education, narrowing the curriculum with its focus on reading and math, and encouraging an overreliance on standardized testing as the sole measure of achievement. By its own measures, NCLB has been largely unsuccessful; while there has been some increase in fourth grade math scores of U.S. students, reading scores have not increased, and no state has reached the required 100% proficiency bar.¹

Under the Obama administration, NCLB has been revised and renamed as the Every Student Succeeds Act (ESSA), with the intention of giving more control to states. While this control is still primarily test-based, the Fordham Institute contends the door is now open to construct approaches toward what they view as the unmet needs of high-scoring students. The ESSA will be fully in place in the 2017-18 school year.

These two closely related reports from the Thomas B. Fordham Institute – High Stakes for High Achievers: State Accountability in the Age of ESSA,² written by Michel J. Petrilli, David Griffith, Brandon L. Wright, and Audrey Kim, and High Stakes for High Schoolers: State Accountability in the Age of ESSA³ – assert that states are not adequately attending to the needs of high-achieving students, and that state accountability systems under ESSA should be adjusted in order to incent districts to address those needs. This is, they argue, not only a question of “fairness,” but also an issue of national importance. “We must also remember,” states the foreword of the first report, “that the country’s future economic competitiveness, scientific leadership, and national security depend on how successfully we maximize the learning of our ablest children.”⁴ Similar in tone to the much-critiqued 1983 A Nation at Risk⁵, these reports link the security and success of the United States to the educational attainment of its highest achieving students.

This review will examine the claims and implications of the reports.
II. Findings and Conclusions of the Reports

Published in two parts, the first focusing on elementary and middle schools and the second focusing on high schools, these reports argue that the No Child Left Behind act was marred by a “pernicious flaw.” This flaw, the authors contend, was that NCLB’s provisions caused schools to focus their attention on raising the achievement of their low-scoring students, while “ignoring the educational needs” of students who scored beyond proficiency. Because of this, they speculate, the United States has seen “smaller gains for its top students” over the past twenty years.

To this end, the reports’ authors propose and apply a rating system to examine and assess how each “state’s current (or planned) accountability systems attend to the needs of high-scoring students (equated with “high-achieving” students by the authors of these reports) and how those systems might be redesigned under the Every Student Succeeds Act (ESSA) to better serve all students.” The authors put forward four steps that states should take to make sure that the needs of these students are prioritized under ESSA. These are: 1) schools should be rewarded for moving more students to an “advanced” level; 2) schools should be rated on whether they use a student growth model that considers the progress of students at all achievement levels, rather than only those below the proficient line; 3) “gifted” or “high-achieving” students should be included as a subgroup, and this group’s results should be reported; 4) “growth for all students” should count for at least half of a school’s summative rating.

The reports turn these qualities into a rating system that they use to assess the accountability system of each state for its attention to “the needs of high achievers.” Using this rating system, the eleven states that do not calculate summative school ratings can receive a maximum of three stars, while the thirty-nine states plus the District of Columbia that do (or intend to) calculate summative school ratings can receive a maximum of four stars.

Employing this system, the reports conclude that only four states – Arkansas, Ohio, Oregon, and South Carolina – have “truly praiseworthy” approaches to incentivizing districts to take measures the authors feel would improve the education of high-achieving students. Five states – Nevada, North Carolina, Ohio, Oregon, and Wyoming – report the results of high-achieving or gifted students as a separate subgroup, and fourteen states and the District of Colombia give (or plan to give) extra credit to schools who reach an advanced level. The authors bemoan draft regulations for the ESSA that may limit states to measuring academic achievement with proficiency rates (how many students reach proficiency) rather than growth measures. The reports’ major recommendation is that the Department of Education “allow states to rate academic achievement using a performance index.”

In addition to this rating system, the high school focused report advocates for rewarding schools with accountability points when students earn college credit through Advanced Placement and International Baccalaureate coursework and early college programs (but only those that confer credits which “transfer to bona fide colleges”).

The reports’ authors argue that the current state of transition from NCLB to ESSA, in which
III. The Reports’ Rationale for Its Findings and Conclusions

These reports contend that the nation’s economic future is at stake if the education of high-achieving students is neglected. As the foreword warns, “We must also remember, though, that the country’s future economic competitiveness, scientific leadership, and national security depend on how successfully we maximize the learning of our ablest children.” Low-income, high-achieving students are at particular risk, the authors explain, because they attend schools “which face the greatest pressure to raise the test scores of their lowest-performing students” and that “typically face a host of other challenges.”

The reports’ authors put forward four criteria for assessing each state’s approach to prioritizing the achievement of high-achieving students, each linked to how states assess and reward schools for gains made by these students on standardized assessments. The reports’ findings and conclusions are then based on these criteria, as derived from publically available data found on state department of education websites and school report cards. The authors contend that, under ESSA, state policymakers should “ensure that their schools have ample incentives to educate those children, and all children, to the max;” suggested incentives are largely based on students’ performance on statewide standardized tests.

IV. The Reports’ Use of Research Literature

There is little research referenced in these reports. Although the reports are based on the idea that high-achieving students have suffered under NCLB, the sources cited for the claim that “students just below the bar were most likely to make large gains in the NCLB era, while high achievers made lesser gains,” do not actually provide clear and convincing evidence to support that assertion. Indeed, one of the sources provided makes the reverse assertion that, “At the extreme high end of the achievement distribution there is also little evidence that students are harmed by NCLB...Rather, these figures suggest the redistribution of gains is from middle high to middle low...However, the magnitude of these effects is small compared to mean grade-level gains.”

No research is cited to explore the complex issues of tracking and ability grouping that underlie the reports’ notions about the special needs and abilities of high-achieving students, despite the existence of a robust body of research on the ways that tracking can exacerbate educational inequalities linked to race and socioeconomic status.

Despite the reports’ aim of drawing attention to the needs of high-achieving students, no
research is cited that describes best practices for this student group or to support the reports’ approach of using high-stakes testing to improve the educational experiences of high achievers.

No research is cited that explores the complexities and potential pitfalls of relying on standardized assessment to both assess learning and incentivize policy change, despite the reports’ reliance on standardized testing as the sole incentive for drawing attention to high-achieving students.17

Although there is much research to support the idea that the persistence of residential and school segregation and resource disparity among schools and communities are at the root of differences in educational outcomes,18 the reports’ authors insist that “there’s no direct link between what goes into a school by way of resources and what comes out by way of student learning.”19

V. Review of the Reports’ Methods

1) Assumptions about how schools educate high-achieving students

These reports are based on assumptions about the education of high-achieving students that are unexamined and unsupported by evidence. At best, these assumptions reveal a shallow understanding of issues of teaching and learning; at worst they appear to be based on stereotypes about particular types of students.

First and foremost is the assumption that not including a subgroup in the state’s high-stakes testing system means that educators are not paying attention to the students in that group. In the foreword to the first report, Chester Finn asserts that, “To put it bluntly, NCLB did some good for America’s struggling pupils, but for high achievers, it mostly just hit the education pause button.”20 In the foreword to the second report, Finn once again raises an alarm, claiming “There are hundreds of thousands of American teenagers ready to work harder, reach higher, and go further, if only we give them the chance. Many are kids of color and come from poor families. They deserve our attention.”21

Secondly, no evidence is provided to support these assertions (see previous section) that educators are not attending to the needs of high-achieving students, and no evidence is provided for the contention that an increased focus on standardized testing will lead to increased gains for this group. The reports do not take into account that high-achieving students, particularly in high school, are enrolled in “advanced placement,” “international baccalaureate,” and/or “dual enrollment” classes that have curricula, standards, metrics and incentives that are designed to encourage achievement. These regular features of schooling, as well as high-achieving students’ desires to build good transcripts for college competitiveness, may be greater incentives for the high performance of this group of students than
high-stakes tests in literacy and mathematics that are given only once during students’ high school careers and have no impact on their future success.

Third, some of the reports’ assertions appear to rest on assumptions derived from stereotypes of low-income students and students of color. In the introduction to the second report, Finn calls for shifting the focus from increasing enrollment in courses, such as AP and IB, that help students earn college credit while in high school, to “rewarding schools where lots of kids pass AP tests.” Rewarding “access” rather than “performance” is a problem, he claims, writing that,

> The impulse to get more students, especially poor and minority youngsters into such advanced options is entirely commendable, but here, too, a worthy goal can have unintended side effects – in this case, by leading to the inclusion of students who aren’t actually prepared to succeed in more challenging academic settings.22

Continuing in this vein, the author bemoans the fate of students who are “pushed through” the doors of higher level classrooms “who may (through no fault of their own) not be up to the challenges within.”23 Providing no evidence to support his implication that high-level classes are filled with students (implicitly “poor and minority”) who are not “actually prepared” for challenging coursework, the author’s assertions are built on stereotypes of unprepared low-income students and students of color rather than actual data.

2) Overstatement of the potential of “growth models” to encourage emphasis on the needs of high-achieving students

As Harvard University educational assessment and testing policy expert Daniel Koretz writes in *Measuring up: What educational testing really tells us*,

> Unrealistic expectations about testing are everywhere. They seem to rest on an inconsistent, even paradoxical view of the complexities of measurement...there seems to be a widespread faith in the wizardry of psychometrics, a tacit belief that no matter what policymakers and educators want a test to do, we can somehow figure out how to make it work.24

These two reports put forward a simplified view of student growth measures and convey an exaggerated sense of what they can reveal. There are several areas in which the reports do not take the complexities of growth measures into account.

First, within the parameters of NCLB/ESSA, measuring growth provides, at best, fairly limited insights into student learning pathways. Under NCLB and ESSA, students are tested in language arts and mathematics in grades 4-8, and once in high school. Because it is most accurate to have three prior measures for an accurate growth measure, and because comparing eighth grade scores with eleventh grade scores is unreliable, growth can really only be accurately ascertained in seventh and eighth grade, and only in two out of all the various school
subjects. This is not much evidence upon which to assess a state’s performance in this area.

Measuring growth on different parts of the scoring spectrum is complex and not parallel; moving a point at one part of the scale is not equivalent to moving a point at a different part of the scale. A Brookings Institute report explains, “Despite common perceptions, a 5-point gain at the bottom of the test score distribution may not mean the same thing in terms of additional knowledge as a 5-point gain at the top of the distribution.” Using one common method for calculating scores, for example, “produces what psychometricians sometimes refer to as “shrunken” estimates of student ability...this approach to scoring will reduce the proportion of students classified into the top and bottom performance categories, as it pushes high and low scores toward the mean.”

For this reason, comparing growth on the low and high ends of the scoring range is more complex than measuring one set of points against another; calculating growth in this way will achieve a false result, and can result in an erroneous judgment about whether a school is making progress with particular groups of students.

High schools “tend to lack key elements needed to track growth,” making it quite difficult to measure growth among the focus population of the second report. Most states only assess students once in high school, making it impossible to ascertain yearly growth for this population. When schools do test annually, those tests are usually by subject (e.g. Geometry; United States History) rather than by grade level, “and are not typically vertically aligned and scaled,” making tracking growth difficult.

Measuring growth using standardized assessments is more complex than these reports would lead us to believe. As Jacobs points out, “While psychometricians (experts in the theory and methodology of psychological measurement) are familiar with these issues, in our experience most others are unaware of them and, as a result, frequently misuse test scores.”

3) The reports do not consider the educational implications for high-achieving students of rewarding the use of more narrow, high-stakes forms of assessment

The two reports assume that incentivizing the use of high-stakes forms of assessment to ascertain the achievement of students on the high end of the achievement spectrum will result in educational improvements for this group. However, there is no evidence to support the idea that an emphasis on accountability testing will result in the high-quality, creative, accelerated learning experiences that are needed by this group of students. Indeed, there is evidence to support the conclusion that overreliance on standardized forms of assessment applied only in math and language arts leads to a narrowing of the curriculum and an emphasis on test preparation, neither of which represent best practices in education for high-achieving – or any other – students.

Moreover, the reports do not consider the uneven distribution of highly qualified teachers, adequate resources and other forms of support that evidence suggests make a real difference for high-achieving students from low-income and historically underrepresented
backgrounds. In so doing, the authors avoid examining some of the causes for the gap in achievement between high-achieving students from white and higher income families and their low-income peers and peers of color, in favor of an approach that promises only to incentivize a test-focused curriculum for all.

4) The reports do not consider the original goals of NCLB and ESSA and the consequences of their recommendations for the role of education in a democracy

By encouraging states to direct resources toward their higher achieving students, these reports fail to consider the core mission of the legislation they seek to utilize. NCLB is an update of the 1965 Elementary and Secondary Education Act (ESEA). Passed in 1965 as part of the Great Society program, its primary mission was “to strengthen and improve educational quality and educational opportunities in the Nation’s elementary and secondary schools,” particularly “children of low-income families” and “disadvantaged children.” NCLB reauthorized this act, continuing the primary focus on “improving the academic achievement of the disadvantaged.” The current ESSA states one of its primary purposes as “to provide all children significant opportunity to receive a fair, equitable, and high-quality education, and to close educational achievement gaps.” No matter one’s stance on the efficacy of the methods employed, all of these Acts were clearly designed with the same purpose – to attempt to address inequalities in education that arise from historical and structural inequalities in U.S. society. By advocating for a shift in focus (and therefore resources) to the testing of high-achieving students, these reports risk the core mission of these Acts.

VI. Review of the Validity of the Findings and Conclusions

These reports raise an important concern – the education of high-achieving students in our nation’s public schools – but fail to engage it productively. By relying solely on standardized testing to create a policy change, the reports incentivize a narrow, testing-focused approach to educational improvement that is not supported by research on teaching and learning. The reports deflect attention from the structural economic inequalities that are the major source of educational disparities, repeating disproven assertions about the lack of impact of resources on educational opportunity and instead encouraging the use of problematic assessment measures to address complex social and economic issues.

VII. Usefulness of the Reports for Guidance of Policy and Practice

In conclusion, these reports can contribute to collective discussion of the nature and role of federal mandates in education, but must be approached with caution. Calling attention to the education of high-achieving students is not in itself a problem. However, the policy
changes recommended by the reports are overly simplistic and do not take into account the complexities of evaluating student growth, the impact of high-stakes testing on the quality of students’ educational experiences, and the needs of the underserved students who are the focus of the ESSA.
Notes and References


30 No Child Left Behind Act, 114 Congress, S. 1177, 2015; Sec. 1001
<table>
<thead>
<tr>
<th><strong>Documents Reviewed:</strong></th>
<th>High Stakes for High Achievers: State Accountability in the Age of ESSA; High Stakes for High Schoolers: State Accountability in the Age of ESSA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors:</strong></td>
<td>Michael J. Petrilli, David Griffith, Brandon L. Wright, &amp; Audrey Kim; Michael J. Petrilli, David Griffith, &amp; Brandon L. Wright</td>
</tr>
<tr>
<td><strong>Publisher/Think Tank:</strong></td>
<td>Thomas B. Fordham Institute</td>
</tr>
<tr>
<td><strong>Document Release Date:</strong></td>
<td>August 2016; November 2016</td>
</tr>
<tr>
<td><strong>Review Date:</strong></td>
<td>January 31, 2017</td>
</tr>
<tr>
<td><strong>Reviewer:</strong></td>
<td>Beth C. Rubin, Rutgers University</td>
</tr>
<tr>
<td><strong>E-Mail Address:</strong></td>
<td><a href="mailto:beth.rubin@gse.rutgers.edu">beth.rubin@gse.rutgers.edu</a></td>
</tr>
<tr>
<td><strong>Phone Number:</strong></td>
<td>(848) 932-0677</td>
</tr>
</tbody>
</table>