

Reconciling Long-term Education Policy Goals with Short-term School Accountability Models

By Jim Soland, Yeow Meng Thum, and Gregory King

KEY FINDINGS

- Schools held accountable for short-term student growth are often very different than ones that would be identified under ESSA if more years of data were used. Rank order of schools varies greatly depending on timespan.
- School accountability, as it currently stands, could be doing a disservice to the schools that are actually doing the most to prepare their students in the long term.

There is a conundrum that underlies much of state and federal education policy: while the aim of many policies is for schools to prepare students for long-term success, like college and career readiness, those same schools are often held accountable for their contributions to student growth over much shorter time periods.

Under the Every Student Succeeds Act (ESSA), a majority of states now include student growth estimates, not just achievement at a single point in time, in accountability measures.

Under ESSA, 47 states plan to use student growth as an accountability indicator; 33 states weight growth the same or more than achievement. Most use current year test scores and data from only one or two years prior to calculate growth.

These measures have consequences for schools, as well as teachers and students: ESSA requires states to identify and intervene in the bottom 5% of schools. Growth measures used by states are often based on student test scores from two to three years of data. However, many school goals under federal and state policy involve improving much longer-term student outcomes, including preparing students for success in college. Despite the increased use of growth measures, little research to date has investigated the implications of this discrepancy for school accountability.

This study begins to close that gap by examining how much rank orderings of schools change when estimates of student growth are made in short- versus long-term timespans. We used MAP Growth assessment data from a cohort of students in grades 2 through 6 in a state that tests students in math and reading in fall and spring of each year, to estimate the contributions of 570 schools to students' growth throughout elementary school. Our research investigated how different the rankings of school effectiveness would be when based on:

- fall to spring growth from one year,
- spring to spring growth over three years, or
- spring to spring growth over all of elementary school.

While growth during elementary schools is not a perfect proxy for a student's college readiness, research shows that longer-term growth is more associated with outcomes like college enrollment than growth from shorter time periods.

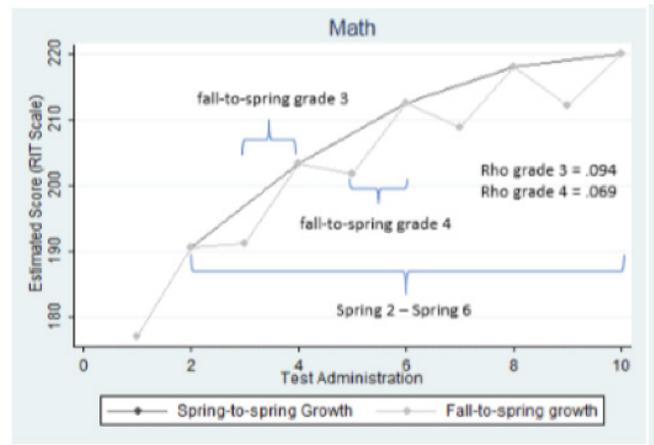
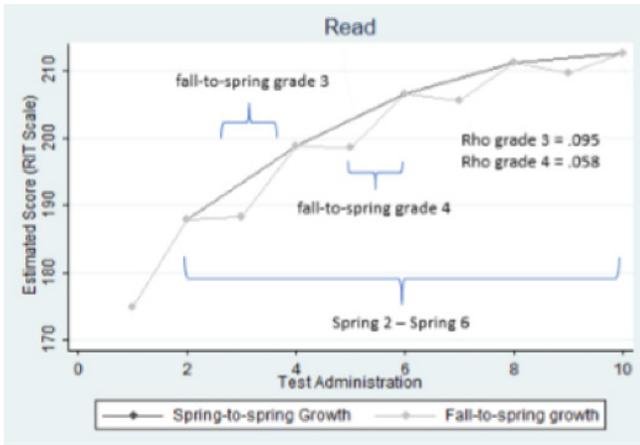
Estimates of school contributions to student growth varied greatly depending on timespan examined

The results showed that there is little relation between estimates of school contributions to growth for a single school year and growth over the course of elementary school. Estimates based on fall-to-spring test score gains from a single year are only correlated with estimates of student growth between spring of second and sixth grade at .10 or below.

School rankings in accountability systems likely differ for short-term versus long-term measures

Even when expanding the model to measure growth over two years (the most common timespan used by states to estimate growth under ESSA), schools identified as low-performing would likely differ from those that would be identified as such if the timespan were expanded to include all of elementary school. More specifically, increasing the timespan to estimate school contributions to student growth over two years improved the correlation of estimates that include grades 2 to 6 (generally, correlations of around .50). Research suggests that determinations made under accountability policies, like those under ESSA, designed to identify extremely low- and high-performing teachers or schools are likely to differ when estimates correlate below .90. Thus, the correlations of .50 were low enough that different schools would likely be identified as low-performing under ESSA when using short-term growth versus growth during the entire span of elementary school, which has been shown to correlate with indicators of college readiness.

In sum, we find that estimated school effectiveness is highly sensitive to timespan, suggesting that short-term accountability policies could produce unintended consequences relative to long-term goals, like preparing students for college.



Compound polynomial model-estimate plot of achievement scores for reading and math by subject and test administration. The model reflects seasonal learning patterns, including the sawtooth in fall-to-spring data representing summer learning loss. Brackets show time spans used in estimates of school contributions to growth.

RECOMMENDATIONS

Ensure policies holding schools accountable for student growth are not at odds with desired long-term student outcomes like college readiness.

By expanding accountability measures to include student growth, states are taking a positive step for long-term student success. Research suggests that holding schools accountable for growth rather than static achievement alone is likely to support efforts around college readiness and that growth is often a better predictor of college readiness than static achievement. The findings of this study suggest that policymakers should engage in a broader conversation about how to hold schools accountable to best support student success and provide the aligned supports for continuous school improvement.

Reexamine and refine how schools are held accountable.

Under federal policy, helping students finish high school, preparing them for college, and closing opportunity gaps that disproportionately affect students of color and economically disadvantaged students are primary aims of the educational system. All of these involve contributions of schools to the long-term growth of their students. This study showed that associations between those long-term estimates and the ones based on shorter-term growth—as typically used under ESSA—differed in statistical and practical significance.

While the current study could not estimate school contributions to student growth through their full kindergarten and 12th grade educations, it revealed a discrepancy between the long-term goals of policy and the short-term nature of accountability that could undermine our best efforts to prepare students for their futures.

As we work together to improve schools and systems, the measures we use to hold schools accountable, and, importantly, how we intervene in schools identified as low-performing, it is important to consider long-term, as well as short-term growth. This may be a useful lever for differentiating support when identifying a school for turnaround.

¹ Koedel, C., & Betts, J. (2010). Value added to what? How a ceiling in the testing instrument influences value-added estimation. *Education*, 5(1), 54–81.
¹¹ Reardon, S. (2016). *School district socioeconomic status, race, and academic achievement*. Retrieved from Stanford Center for Education Policy Analysis: <http://cepa.stanford.edu/content/school-district-socioeconomic-status-race-and-academic-achievement>
¹¹¹ Thum, Y. M., & Bhattacharya, S. K. (2001). Detecting a change in school performance: A Bayesian analysis for a multilevel join point problem. *Journal of Educational and Behavioral Statistics*, 26(4), 443–468.

ABOUT THE AUTHORS

Dr. Jim Soland is a Senior Research Scientist at the Collaborative for Student Growth at NWEA, and is an Adjunct Professor of Quantitative Methodology in the College of Education at Oregon State University. His research focuses on assessment and evaluation policy and practice, with particular emphasis on measuring social emotional learning, test engagement, and estimating teacher and school effectiveness. Soland completed a doctorate in Educational Psychology at Stanford University with a concentration in measurement and policy.



Dr. Yeow Meng Thum is a Senior Research Fellow at the Collaborative for Student Growth at NWEA. His work focuses on methodological research on multivariate, multilevel models for behavioral and educational data, as well as topics in latent variable modeling, causal inference, and the psychometrics of scaling and norming. Thum holds a PhD from the University of Chicago.



Gregory King is a Research Scientist at the Center for School and Student Growth at NWEA. His research focuses on college and career readiness, policies that impact education access and success, as well as data visualizations for the use of reporting information to educators, administrators, and policymakers. King holds a PhD in Educational Research and Policy Analysis from North Carolina State University.



ABOUT THE COLLABORATIVE FOR STUDENT GROWTH

The Collaborative for Student Growth at NWEA is devoted to transforming education research through advancements in assessment, growth measurement, and the availability of longitudinal data. The work of our researchers spans a range of educational measurement and policy issues including achievement gaps, assessment engagement, social-emotional learning, and innovations in how we measure student learning. Core to our mission is partnering with researchers from universities, think tanks, grant-funding agencies, and other stakeholders to expand the insights drawn from our student growth database—one of the most extensive in the world.



NWEA is a not-for-profit organization that supports students and educators worldwide by providing assessment solutions, insightful reports, professional learning offerings, and research services. Visit [NWEA.org](https://www.nwea.org) to find out how NWEA can partner with you to help all kids learn.

© NWEA 2019. MAP and NWEA are registered trademarks, and MAP Growth is a trademark, of NWEA in the US and in other countries.

MAR19 | KAP3490