

Don't Forget Curriculum

By Grover "Russ" Whitehurst

There is a broad consensus across the political spectrum on the need for K-12 education reform. President Obama has committed himself to "reform America's public schools." Likewise, George W. Bush began his administration by calling for "real reform" in education.

"You do what works for the kids."

-U.S. President Barack Obama

There is also surprising agreement on the means of reform. With the Obama administration focusing on early childhood programs, common standards, charter schools, and more effective teachers, no one scoffed when Bush Secretary of Education Margaret Spellings said of Obama, "He is saying a lot of things that sound all too familiar to me. I want to sing right along."

However, the two administrations diverge with respect to the place of curriculum in education reform. By curriculum I mean the content and sequence of the experiences that are intended to be delivered to students in formal course work. Curriculum includes teaching materials such as those that can be found in commercial textbooks and software applications. It also includes the pedagogy for delivering those materials when teachers receive guidance on how to teach the curriculum, or when software manages the pacing, prompts, and feedback that students receive as they engage with the materials. Of course, the same curriculum can often be delivered with different instructional strategies.

The nation has made a large investment in education R&D to create and evaluate curriculum materials. Content experts and teachers think it is critically important. States with adoption policies for textbooks agree. The What Works Clearinghouse, established during the Bush administration, directed much of its effort to summarizing curriculum evaluations for use by practitioners and policymakers. Similarly, the National Center for Education Evaluation within the U.S. Department of Education launched a significant number of rigorous evaluations of curriculum during the Bush years.

The Obama Education Department, in contrast, has been mum on curriculum, with the exception, greeted with consternation by some, of providing guidance on how teachers might incorporate the President's back-to-school TV address into their lesson plans.

One reason sometimes cited for the federal policy makers to shy away from curriculum is the Department of Education Organization Act of 1979, which prohibits the Department from endorsing or sanctioning any curriculum designed to be used in an elementary school or secondary school. Similar prohibitions have been included as boilerplate language in the Elementary and Secondary Education Act, the Individuals with

Disabilities Education Act, and the Education Sciences Reform Act.

Whatever the legislative intent in these prohibitions, the Department has always distinguished between providing information on curricula vs. endorsing it, and between requiring that education agencies that are recipients of federal funds use the best evidence in their choice of curriculum vs. providing a list of sanctioned curricula. Whether it be the What Works guides produced during the Reagan administration, the National Diffusion Network in place during the G. H. W. Bush presidency, the Exemplary and Promising Practices Panels in the Clinton administration, or the What Works Clearinghouse under G. W. Bush, federal executive branch attempts to describe what is working best and to review the research evidence on the effectiveness of particular curricula have always been seen as within the law.

The 2001 reauthorization of the Elementary and Secondary Education Act (No Child Left Behind) pushed the Department of Education close to and in some cases over the line in endorsing curriculum. Well on the safe side of the curriculum prohibition, the Act required that states and local education agencies use scientific research to select curricula for their Reading First program. But it also required that states submit their plans for selection of curriculum materials to the Department for approval. The Department pushed hard to get state plans that favored the curriculum it wanted, and it went so far as to intervene politically when the New York City public schools selected a curriculum not to its liking. States, districts, and frustrated curriculum developers complained that the Department had an implicit list of approved reading programs that was not warranted in legislation or regulation, and was denying approvals to states who did not adhere to it. Congress had hearings, the press had a field day, and heads in the Department rolled.

In light of the legislative prohibitions on endorsing curricula and the political taint surrounding Reading First, one can imagine high-level meetings in the Obama administration in which curriculum and third rail were mentioned in the same sentence. But one can also imagine an administration that is staffed with policy makers who cut their teeth on policy reforms in the areas of school governance and management rather than classroom practice, people who may be oblivious to curriculum for the same reason that Bedouin don't think much about water skiing.

Governance vs. Curriculum Reformers

People who are trying to create more charter schools, or pressure unions to allow more flexibility in hiring and firing teachers, or transform schools into one-stop shops for community needs, do not sort with people who are trying to improve the teaching of fractions or children's reading comprehension. The disciplinary training, job experience, professional networks, and intuitions about what is important hardly overlap between governance and curriculum reformers. For the governance types, teaching resolves to the question of how to get more qualified teachers into the classroom, e.g., "How can we remove the artificial barriers to entry into the profession so that smart people who want to teach don't have to jump through the hoops of traditional teacher training and certification?" For the curriculum reformer, teaching is about specific interactions between students and their curriculum materials as shaped by teachers. For a curriculum reformer, teachers with higher IQs and better liberal arts educations are desirable, to be sure. But just as people with musical talent have to work hard to develop musical skills and have available to them exceptional compositions if they are to be successful musical performers, so too bright aspiring teachers have to learn a lot about how to teach and have good curriculum materials if they are to be effective with students. Thus being smart is the starting point of becoming a good teacher for a curriculum reformer whereas it is often the end point of governance reforms.

Let's assume the Obama administration has ignored curriculum inadvertently because it is staffed with governance people who are simply valuing what they know. If so, then the administration would do well to

heed Obama's assertion that, "you do what works for the kids." The administration should be open to all the categories of reform and innovation that could have an appreciable impact on student learning.

Based entirely on a what works metric, where would curriculum stand with respect to prescriptions for reform endorsed by the administration, including differential pay for teachers and other mechanisms for reconstitute the teaching workforce, charter schools, early childhood programs, and common standards?

Effect Size

Let's consider these policy levers in turn, using something called effect size as a metric. Effect size is a way of representing in numerical terms the strength of a relationship between an educational influence and a student outcome. For example, one could specify the strength of the relationship between parental socioeconomic status and high school graduation rates, or between attending a charter vs. a regular public school and achievement in mathematics. The more familiar test of statistical significance answers the question of whether an association is due to chance. Effect size, in contrast, addresses the strength of the association. Weak and unimportant effects can be shown to be statistically significant with a sufficiently large sample, which is why statistical significance alone is a poor guide to the educational significance of a program.

The simplest way of reporting effect size is as the difference between the means of the two groups on an outcome of interest. For example, the effect of Career Academies, a school-based vocational education program, can be expressed as the \$212 difference between the future monthly earnings of students who were randomly assigned to participate in Career Academies versus those who were not. The same effect can be standardized by dividing by it the standard deviation of the measure on which the difference between the groups was calculated. In this case the standard deviation of monthly earnings for the sample being studied was \$706. This gives an effect size of $\$212/\$706 = 0.30$. A tremendous advantage of a standardized effect size is that it allows comparisons of the strength of effects across different domains and different outcomes. Thus one could compare the 0.30 effect size of Career Academies on monthly earnings with its 0.00 effect size on high school graduation.

Effect sizes can also be computed for the strength of associations between dimensional variables. For example, one can compute the effect size for the relationship between teacher quality and student mathematics achievement as the square of the simple correlation of these two measures. Note that in this case, unlike the previous example of Career Academies, outcomes aren't being compared for discrete conditions to which students are assigned, ideally randomly. Rather the measure of effect size is derived from the degree of association between measures of teacher quality and student achievement

In making sense of effect sizes, it is important to ask the degree to which the design of the study that generated the effect size supports a causal interpretation. That one can calculate a healthy effect size for the relationship between school achievement and whether students attend suburban vs. urban schools does not mean that suburban schools are necessarily better (children attending suburban schools are, on average, from more affluent, educated families than those attending urban schools). Likewise, the strong negative effect size that can be calculated for the relationship between the receipt of a subsidized school lunch and student achievement does not mean that giving kids lunch makes them learn less. That is why social scientists place such a strong premium on research designs that disentangle the effects of correlated variables, with the randomized trial typically being the strongest for this purpose.

Making sense of effect sizes also requires asking whether the effect size is derived from the contrast between groups that have or have not experienced an intervention vs. from an association between variables that may

not map onto any ready policy levers. For example, a measure of the association between teacher quality and mathematics achievement may identify a strong influence of teachers on students, but does not test an intervention or policy to change teacher quality so as to improve student outcomes.

With that background on effect sizes, let's examine what they tell us about the known and likely effects of popular education policy prescriptions the Obama administration is currently looking to fund and push from the federal level including more charter schools, reconstituting the teacher workforce, more early childhood education, and national content standards.

Charter Schools

The Obama administration has made a strong push to expand charter schools by making states that don't permit charters or that cap them ineligible for \$4.35 billion in Race to the Top funds. Both President Obama and Secretary Duncan have also used their bully pulpits to press the case for a reform model in which poor performing traditional public schools are shuttered and replaced by charter schools.

What does research say about the size of the effect of charter schools on academic outcomes when the effect is measured as the difference between performance of students in charter schools and comparable students in traditional public schools?

Studies that have employed large samples of charter schools and controlled statistically for background differences between students, generally find very small differences in student achievement between the two types of public schools. For example, a recent [large study](#) of charters in several states found that they produced achievement gains for students who transferred into them that were, on average, the same as those of traditional public schools in five of the studied locales, and somewhat worse in two. In another [recent study](#) of charter schools in Milwaukee, initial differences favoring charters in mathematics achievement faded out over time. In a [2005 study](#) utilizing test results from the National Assessment of Educational Progress, white, black, and Hispanic fourth graders in charter schools performed equivalently to fourth-graders with similar racial/ethnic backgrounds in traditional public schools.

More positive findings for charters emerge from studies of oversubscribed charter schools in which lotteries were used to determine admission. For example, a [recent study](#) of performance of students in charter schools in New York City found positive effects, with the average effect size for a year spent in a charter school = 0.09 for mathematics and 0.06 for reading. Similar studies with similar results have been reported for oversubscribed charters in Chicago and Boston.

Thus the effect of a typical charter school on student outcomes is not likely to be different from that of a typical traditional public school, but popular, oversubscribed charter schools operating in some large urban school districts have positive effects.

Reconstituting the Teacher Workforce

Although the contribution of teachers to student learning may seem intuitively obvious, a generation of research dating from a 1966 study by sociologist James Coleman had supported the conclusion that nearly all of the differences in achievement outcomes among students were attributable to differences in family background. In contrast, we now know that substantial differences in the achievement of students are attributable to differences in teachers. For example, in a [2004 study](#) that took advantage of the random assignment of students to teachers in the Tennessee class size experiment researchers found that differences in teachers accounted for 0.12 to 0.14 of students' annual mathematics achievement gains.

Note that this is an association-based measure of effect size. It does not represent the effect of any identified intervention or policy. Rather it represents the amount of variability in student outcomes than can be attributed to differences among teachers. It is one thing to conclude, as did researchers in a related [study](#), that, "the difference between being assigned a 25th or a 75th percentile teacher would imply that the average student would improve about one-quarter of a standard deviation relative to similar students in a single year." It is another thing entirely to construct a policy that would reconstitute the teacher workforce so that 100% of teachers would be of the quality of the top 25% today.

Easing barriers into the teaching profession for academically highly qualified entrants is one policy that has been actively explored for reconstituting the workforce. Teach for America is the best known example of a variety of flourishing alternative pathways into teaching. A [well conducted trial](#) comparing students randomly assigned to TFA teachers vs. other teachers found a 0.15 effect size for math achievement and no effect on reading achievement.

Merit pay is another policy that is being pursued to reconstitute the teacher work force. The conceptual model, based on extensive research on other labor markets, is that paying higher salaries to more effective teachers would keep them in the profession at higher rates than would be the case for teachers not receiving merit raises, and would encourage more talented individuals to enter the profession. Over time, this would shift the mean of the quality curve for teachers upward.

There is a small but growing body of research on the effect on student academic outcomes of incentives to teachers to raise student performance. The National Mathematics Advisory Panel issued a report in 2008 that included a systematic [review](#) of research on performance pay for teachers. Across the 14 studies reviewed, all but one found some positive effects on student achievement. The strongest [study](#) methodologically was conducted in India and involved substantial teacher bonuses for raising student scores. The impact of the incentives at the end of one year was an effect size of 0.15 across reading and mathematics. The degree to which the incentive system employed in India would generalize to the U.S. is unknown.

Early Childhood Programs

The research literature on the effectiveness of early childhood programs is large, spans a half century, and varies dramatically in methodological quality, years of follow-up, and outcome measures. This makes it impossible to summarize this body of research responsibly in a few paragraphs. Advocates for increased investment in preschool programs trumpet the findings from two small demonstration programs implemented in the 1960s and 1970s, the Perry Preschool Program in Michigan and the Abecedarian Program in North Carolina. These programs produced long-term effects on social outcomes such as employment. The Abecedarian project also had impacts on reading achievement of approximately 0.45 when participants were 15 years old.

Evidence on the impact of current large-scale federal preschool initiatives is less impressive and incomplete. A recent large randomized [trial](#) of Head Start found an effect size of 0.24 on letter naming at the end of the Head Start year, but no measurable effects on vocabulary. The children in this federal study were followed-up when they were in the 3rd grade in 2007 and 2008, but the Department of Health and Human Services has neither released the findings from that follow-up nor announced a timetable for release. Another federally funded preschool program, Even Start, has shown no effects on child outcomes in two separate rigorous [evaluations](#). A frequently touted program that is slated for expansion by the Obama administration, the Nurse-Family Partnership, has been subjected to only one rigorous [evaluation](#) in which the academic achievement of children was examined as they progressed through school. There was an effect size of 0.09 (n.s.) for reading and mathematics test scores for grades 1 – 3.

There is enough evidence from an older generation of demonstration programs and sound reasons from our knowledge of developmental science to invest in high quality early intervention programs for children who are unlikely to receive the developmental support they need at home. At the same time, we need to acknowledge that the evidence for substantive impact on children of current versions of federally supported preschool programs is weak. The older generation of demonstration programs that had large impacts were very expensive. Another benefit of effect sizes is to allow for comparative benefit-cost analyses. Scarce resources need to be allocated to get the biggest bang for the buck.

Content Standards

The whole standards and accountability movement is grounded in the assumption that high quality content standards for what students should know and be able to do are essential elements of reform. In keeping with this assumption, the Common Core State Standards Initiative, a joint effort by the National Governors Association Center for Best Practices and the Council of Chief State School Officers, has signed up 48 states and 3 territories to develop a common core of state standards in English-language arts and mathematics for grades K-12. Secretary Duncan has praised the effort, made participation in it a prerequisite for Race to the Top Funding, and set aside \$350 million in stimulus act funding to develop state assessments aligned to the whatever common core standards emerge.

What do we know empirically about the association between the quality of content standards and student achievement? Finding very little research on this issue, I and my Brown Center colleague, Michelle Croft, conducted an exploratory analysis of the associations between student achievement outcomes in mathematics at the state level and ratings of state content standards in mathematics conducted separately by the Fordham Foundation and the American Federation of Teachers. The measures of student achievement included 1) fourth grade NAEP mathematics scores for white and black students for the years 2000, 2003, 2005, and 2007 and 2) state gains on NAEP fourth grade mathematics for white and black students from 2000 to 2007 and from 2003 to 2007. The measures of the quality of state standards were the Fordham Foundation ratings of state elementary school math standards from 2005 and the AFT ratings of state elementary school math standards from 2008.

Whether examining results for white or black students for each year of NAEP outcomes or for gain scores over either time period, we found no statistically significant association and very small effect sizes. The largest positive relationship in absolute terms was between Fordham Foundation's rankings and the NAEP gains for white students from 2000 to 2007, with 0.035 of the differences in state gains in achievement accounted for by the quality of state standards. Many of the associations with scores in a particular year were negative, e.g., the Fordham rating of state math standards in 2005 was negatively correlated with state performance on NAEP for both white and black students in the same year. Again, all of the effect sizes were small and none

of the associations was statistically significant.

The lack of a systematic relationship is illustrated when reviewing the data for the “high” standards and the “low” standards states. Massachusetts, for instance, has high standards according to both the Fordham Foundation and the AFT and high NAEP scores. However, New Jersey has low quality content standards on both the Fordham Foundation and on the AFT scales, but scores comparably to Massachusetts on NAEP. Likewise, for gains in NAEP scores from 2000 to 2007, there is no systematic relationship between the “high” standards and the “low” standards states. California is given the highest Fordham Foundation rank and has high gains in NAEP scores. Arkansas, which receives a very low Fordham Foundation rank, has almost identical gains to California on NAEP from 2000 to 2007.

These results correspond closely to those from a 2008 [study](#) by the National Center For Education Statistics that found substantial differences among states in how much mathematics students needed to know to be deemed proficient on the state’s own assessment, but virtually no relation between the level at which a state set its standard and the mathematics achievement of its students.

The absence of a correlation between ratings of the quality of standards and student achievement and between the difficulty of state standards and student achievement raises the possibility that better and more rigorous content standards do not lead to higher achievement – perhaps standards are such a leaky bucket with respect to classroom instruction that any potential relationship dissipates before it can be manifest. Alternatively, the Fordham and AFT ratings may not have captured the qualities of state content standards that drive achievement. Or NAEP may be too blunt an instrument to detect the influence of the quality and rigor of standards.

The lack of evidence that better content standards enhance student achievement is remarkable given the level of investment in this policy and high hopes attached to it. There is a rational argument to be made for good content standards being a precondition for other desirable reforms, but it is currently just that – an argument.

What About Curriculum?

There are two methodologically credible approaches to estimating effect sizes for curriculum. The first, and by far most frequently used, is to compare outcomes for students receiving a branded or new curriculum with outcomes for a similar group of students who experience “business as usual.” Business as usual is typically an unbranded mix of whatever teachers and schools were employing before the new curriculum was introduced. If classrooms have been randomly assigned to continue what they have been doing or to implement the new curriculum, the difference in outcomes for students in the two groups of classrooms is the measure of the effect, or impact, of the new curriculum.

A significant problem with this method is that the nature of the business as usual, the so-called counterfactual, is both not specified and likely to differ from study to study as research is carried out in different schools and locales. Further, the outcomes that are measured in a study of one curriculum may be different than those used in a study of another curriculum in the same domain. This means that the relative effectiveness of curriculum interventions cannot be directly inferred by comparing their effect sizes. Suppose curriculum A is tested in preschools in which business as usual includes an introduction to numbers and counting, whereas curriculum B is tested in preschools in which math instruction is largely absent. The resulting effect size on a mathematics achievement assessment is 0.10 for curriculum A and 0.30 for curriculum B. This might be taken to mean that curriculum B is better, but since the effect size is just the difference in performance between the two conditions, it might also reflect the weaker competition faced by curriculum B. Just as it is difficult to

compare race horses who have never performed on the same tracks with similar levels of competition, so too is it difficult to compare curriculum that have not been tried out side by side with the same outcome measures.

The second method for estimating curriculum effects does just that. In a comparative effectiveness trial, students or classrooms or schools are randomly assigned to receive one of two or more curricula, and the same outcomes are measured for all participants. Since everything about the students (or classrooms or schools) except the differences in curriculum is averaged out through random assignment curriculum effect sizes can be compared directly.

A recent comparative effectiveness [trial](#) of four elementary school math curricula carried out by the Institute of Education Sciences (IES) within the Department of Education demonstrates the power of curriculum as a policy lever for education reform. Just seven math curricula constitute 91 percent of the curricula used by K-2 educators. Should these curricula differ substantially in effectiveness, the implications for policy and practice would be significant.

The IES study involved randomly assigned schools in each of four participating school districts to four curricula. The relative effects of the curricula were calculated by comparing math achievement of students in the four curriculum groups. The curricula were selected from among those that dominate the market, with the goal of having as much diversity in curriculum approach as possible.

Students were pretested in the fall and post-tested in the spring of first grade on a standardized assessment of mathematics. Two of the curricula were clear winners. The spring math achievement scores of Math Expressions and Saxon Math students were 0.30 standard deviations higher than for students experiencing Investigations in Number, Data, and Space, and 0.24 standard deviations higher than for students experiencing Scott Foresman-Addison Wesley Mathematics. This means that a student's percentile rank would be 9 to 12 points higher at the end of a school year if the school used Math Expressions or Saxon, instead of the less effective curricula.

Even larger curriculum effects have been shown repeatedly in business as usual trials. For example, in an Institute of Education Sciences [study](#) of 14 preschool curricula, each the subject of a separate trial against a business as usual condition, a curriculum consisting of DLM Early Childhood Express supplemented with Open Court Reading Pre-K generated very positive effects on a number of child academic outcomes measured on follow-up at the end of kindergarten. For example, the effect size on a test of early reading ability was 0.76; it was 0.48 for a measure of vocabulary. It is very hard to move the vocabulary scores of young children from low-income families (note the previously mentioned null finding for Head Start). This particular preschool curriculum effect is equivalent to moving students 18 percentile points higher on vocabulary in kindergarten.

A cursory exploration of topic reports from the [What Works Clearinghouse](#) will reveal many curriculum and program interventions with very large effect sizes for important outcomes. For example, two programs designed to reduce dropping out, Accelerated Middle Schools and Check and Connect, produce effect sizes of approximately 1.00 for progressing in school. Several beginning reading programs generate effect sizes for alphabets (sound-letter correspondence) of above 0.80. These are large impacts in anyone's categorization scheme.

Curriculum vs. Other Policy Levers

Summary of Effect Sizes

<i>Charters</i>	
Charter schools in general	0.00 mathematics
Oversubscribed NYC charter schools	0.09 mathematics
<i>Reconstituting the teacher workforce</i>	
Merit pay for teachers in India	0.15 reading and mathematics
Teach for America	0.15 mathematics
<i>Preschool programs</i>	
Abecedarian Preschool	0.45 reading
Head Start	0.24 letter naming
Head Start	0.00 vocabulary
Even Start	0.00 vocabulary
Nurse Practitioner Partnership	0.09 reading & math test scores
<i>State standards</i>	0.00 mathematics
<i>Curriculum comparisons</i>	
More effective math curricula	0.30 mathematics
Most effective preschool curricula	0.48 vocabulary
Most effective dropout preventions	1.00 progressing in school
Most effective early reading programs	0.80 alphabetics

Anyone interested in, “doing what works for the kids,” should pay attention to this table. The particular point estimates of effect sizes are not important, but the general range and magnitude of effects is. Curriculum effects are large compared to most popular policy levers.

Further, in many cases they are a free good. That is, there are minimal differences between the costs of purchase and implementation of more vs. less effective curricula. In contrast, the other policy levers reviewed here range from very to extremely expensive and often carry with them significant political challenges, e.g., union opposition to merit pay for teachers. This is not to say that curriculum reforms should be pursued instead of efforts to create more choice and competition through charters, or to reconstitute the teacher workforce towards higher levels of effectiveness, or to establish high quality, intensive, and targeted preschool programs, all of which have evidence of effectiveness. It is to say that leaving curriculum reform off the table or giving it a very small place makes no sense. Let’s do what works for the kids, and let’s give particular attention to efficient and practical ways of doing so.

What Should Happen Next?

As the above data illustrate, we conclude that the effect sizes for curriculum are larger, more certain, and less expensive than for the Obama-favored policy levers. We recommend that the Obama administration undertake the following actions to better integrate curriculum innovation and reform into its policy framework.

The Department of Education, through the Institute of Education Sciences, should fund many more comparative effectiveness trials of curricula and other interventions, both through its National Center for Education Evaluation and through competitive grants to university-based researchers. The Obama

administration has clearly recognized the importance of comparative effectiveness research in health care reform. It is no less important in education reform.

The Department of Education's comprehensive technical assistance centers and other technical assistance contractors should prioritize the dissemination of information from comparative effectiveness trials.

The Institute of Education Science's What Works Clearinghouse should give additional weight to findings from comparative trials in summarizing the results of studies. But since it will be years before there is enough stock of comparative trials to guide most education decisions, the What Works Clearinghouse should do more to support comparisons of the effectiveness of curricula that have been studied using business as usual counterfactuals. In particular, it should develop and publish tables of expected annual gains for widely used assessment instruments and report gain scores for the intervention condition in each study of curriculum on which it reports findings, preferably in an easily understood metric such as months of schooling. This would allow consumers to obtain a sense of how various curricula perform with respect to national norms for growth. In the absence of true comparative data on curricula under consideration for adoption, the wisest choice in many circumstances would be the curriculum that has been found to generate decent effect sizes in credible trials against business as usual AND the largest growth with respect to national norms.

The reauthorization of the Elementary and Secondary Education Act should include a provision requiring state and local agencies to review research on curriculum effectiveness and justify curriculum decisions in light of that evidence for purchases over a certain fixed dollar amount, or for purchases intended for use district wide or statewide, regardless of cost. This review and justification should be available for public inspection. The public reporting and the scrutiny it is likely to engender is an alternative to the type of federal curriculum oversight found in Reading First. It incentivizes evidence-based education while allowing states and local education agencies the flexibility they need to respond to local circumstances, including their own data.

Author

Grover “Russ” Whitehurst is a senior fellow and director of the Brown Center on Education Policy at the Brookings Institution. Previously, he was the director of the Institute of Education Sciences in the U.S. Department of Education.

About the Brown Center on Education Policy

Established in 1992, the Brown Center on Education Policy conducts research and provides policy recommendations on topics in American education. The Brown Center is part of The Brookings Institution, a private nonprofit organization devoted to independent research and innovative policy solutions. For more than 90 years, Brookings has analyzed current and emerging issues and produced new ideas that matter - for the nation and the world.

About the Brown Center Letters on Education

The Brown Center Letters on Education will be published several times each year. Each Letter will address an issue that is of current interest to education policy makers or practitioners by providing background, reporting relevant evidence, and making recommendations for action. Authors of Letters will be scholars with experience bridging the worlds of research and practice. Their charge is to make specific and practical recommendations that are grounded in a careful and dispassionate consideration of evidence. While the recommendations are to be grounded in evidence, they will rarely be compelled by it. Other scholars looking at the same evidence could generate different interpretations.

Brown Center on Education Policy

The Brookings Institution
1775 Massachusetts Ave. NW
Washington DC, 20036
202.797.6090
202.797.6144 (f)
www.brookings.edu/brown.aspx

This publication was made possible in part by a grant from the W.T. Grant Foundation. The statements made and views expressed are solely the responsibility of the author(s).