

**MAKING RESEARCH IN
EDUCATION FINANCE
AND POLICY MATTER NOW**

Amy Ellen Schwartz

Institute for Education
and Social Policy
New York University
665 Broadway, 8th Floor
New York, NY 10012
amy.schwartz@nyu.edu

INTRODUCTION

Research in education finance and policy has flourished over the past twenty years as No Child Left Behind (NCLB) and a wide range of school reform efforts spurred demand for scientific evidence identifying “what works.” Research funding has been generous, buoyed by both favorable economic conditions and the sense that research will provide solutions to persistent problems in American schooling. It has been a good time for education research.

The good news is that we have made significant strides over the past decades. The quality and availability of data have improved substantially, reflecting improvements in administrative data from school districts and states (e.g., Texas, New York, and Florida) and survey data from the U.S. Department of Education containing longitudinal data on both students (e.g., National Education Longitudinal Study, Early Childhood Longitudinal Studies) and schools and school districts (e.g., the Common Core of Data). Perhaps equally important has been the development and spread of sophisticated research methodologies that can be used to identify causal relationships between educational interventions and outcomes and, more generally, to disentangle the causes and consequences of student achievement. Although not entirely new, natural experiments, instrumental variables, regression discontinuity designs, hierarchical linear models, and randomized controlled trials are now fixtures in education policy research.

Unfortunately, our successful use of and enthusiasm for new techniques and data have not been matched by comparable success in identifying solutions to pressing problems and resolving continuing disputes about efficacy and efficiency. Indeed, the results have been fairly modest—yielding more insight into “what doesn’t work” than what does. It is not surprising that policy makers and educational leaders find this harvest disappointing.¹ As research funding from foundations and governments tightens with the economic downturn and a new presidential administration takes the helm, the time is ripe to reevaluate and consider how to make research in education policy and finance matter.

In the end, the key to making education policy research matter is asking questions that matter—about pressing problems that affect large numbers of students in a broad range of circumstances—and providing useful answers and solutions that are feasible, practical, and implementable under realistic circumstances. Why do some students succeed while others do not? What can and what should the public sector do about it? These are the fundamental questions.

I see three key implications for making education policy research matter now. First, methodology must follow from the question (rather than vice versa). This means we must look beyond the methodologically neat and fashionable and focus on bringing to bear the theoretical and methodological tools necessary and appropriate to answering the questions that matter. This may mean more descriptive work that measures, documents, and locates problems. It may mean fewer randomized controlled experiments that yield internally valid estimates of programs implemented in a small number of schools or classrooms and more emphasis on analyses of the impact of policies and programs implemented and practiced by states, school districts, and schools. Further, we need to think about the theoretical underpinnings of our empirical work and bring theory to bear in both design and interpretation. Understanding the behavioral responses of parents, teachers, students, and taxpayers is critical to drawing the link between impact estimates and policy changes.

Second, we need to look outside the traditional boundaries of education research to understand how nutrition, health care, housing, and other factors affect student outcomes. We need to look beyond the school day and the current academic year to understand how and when the past shapes the present and how out-of-school time complements or complicates school experiences. While few would argue that academic outcomes depend only on what happens in the classroom, during the school day, and during the current academic year, much of our empirical research proceeds as if it were so. Understanding the

1. See, for example, Glod 2008 and Viadero 2009.

contributions of outside-the-classroom factors, out-of-school time, and prior year experiences is critical to understanding how schools shape and improve student outcomes.

Finally, we need research on the large, diverse population of urban school children, the challenges posed by poverty, immigration, mobility, race, and ethnicity, and the systemic challenges of large urban school districts. While there are challenges facing students and schools across the country, those facing cities and their students seem to me to be particularly compelling. I elaborate on these below.

METHODOLOGY MATTERS . . . THEORY TOO

Good research provides answers that are useful, informative, and, ultimately, right. Unfortunately, the incentives for researchers are poorly aligned with these objectives. Research projects are driven by considerations about what is fundable and publishable in prestigious journals. These incentives easily translate into an emphasis on the clever over the insightful, on the methodologically fashionable over the useful, and on quickly executed studies with straightforward results over sustained studies yielding full, nuanced, robust answers. We also need to do better at matching the answer to the question. To some degree, the recent focus on estimating “effects” has eclipsed concerns about *what the estimated effect means*. We can bridge this gap by bringing theory from economics, politics, sociology, and psychology to bear both in the modeling and estimation stages and in the interpretation.

One good example can be found in the research on class size. Over the last decades, a substantial body of research has focused on estimating the impact of class size on student performance. Some of this research uses experimental methods—most notably the randomized controlled trial of the Tennessee STAR (Student Teacher Achievement Ratio) experiment. Others use quasi-experimental methods (such as Angrist and Lavy 1999). Still others proceed by modeling the production function for education and using econometric methods to address potential biases.² In large part, the analyses are designed to estimate the impact of changing class size on the performance of the students in those classes. This is not the same as the impact of changing class sizes in *all schools in a district* or *all schools in a state*. Knowing whether reducing class size in a subset of schools or classrooms has a positive effect on student performance is interesting and important in its own right, but it does not answer the policy question: what is the impact of a statewide policy to reduce class size?

2. Hanushek 1986 is a classic; see also Hanushek 1999 and Nye, Hedges, and Konstantopoulos 1999. The *Handbook of Research in Education Finance and Policy*, edited by Ladd and Fiske (2007), is a useful reference.

Large-scale policy changes involve a host of issues that a randomized, controlled trial will not address, such as induced demand for teachers and classrooms. If existing school buildings are at capacity, reducing class size might reduce school size and induce a demand for more school buildings, which may be of different quality (or character) than the existing schools, with unknown effects on student performance. Reducing class size might affect the composition of schools because it offers more opportunity for segregation and isolation. Of course, implementing a statewide class size reduction policy may create new demand for teachers and lead to a redistribution of teachers across schools, with considerable consequences for students and schools. (The California experience with class size reduction is instructive.) In the end, the finding that reducing class size on a micro scale improves academic performance of the students in those classes is neither necessary nor sufficient evidence for a positive impact of class size reduction on a macro scale. Instead, we need to assess the general equilibrium effects—the impacts on input demand and the mobility of students, taxpayers, and teachers—which requires an entirely different strategy than replicating an experiment to assess the generalizability of the results to different populations or settings.³

The point is that we need to carefully align the design and method with the question and the theory. If the intention is to estimate a local average treatment effect of an intervention implemented at a small scale—shedding light on the science of learning and potential directions for school reforms—then a randomized controlled trial will do well. If the purpose is to uncover systemic effects of large-scale reform—which I would argue is the critical policy question—then we need more than a good clean estimate of a program impact.

Disentangling causal impacts requires theoretical and not just methodological sophistication. As an example, consider the estimation of an education production function—the workhorse of education research—estimated to identify the impact of school resources on student outcomes. Typically, a test score (or gain) is linked to variables capturing school resources and student characteristics (including a lagged score if needed to create a value-added measure).

Imagine that resources are allocated via a formula, which gives money to schools based on the number and characteristics of their students (poverty, special needs, limited English proficient (LEP), etc.) and some institutional features (grade span, building features, etc.). Some of these resources are allocated by unit (e.g., number of teachers) and others by dollars (e.g., book allowance). Notice that the production function regression model includes many, if not all, of the factors that drive the differences in resources. The

3. The work of Tom Nechyba (2000, 2003a, 2003b) provides good examples.

reasons might be precisely the same as the reason these factors are included in the allocation formula: they are likely to capture differences in inputs required to achieve a given output level. The implication is that the estimated coefficients on resources capture the impact of resources not allocated via this formula—that is, not given because of differences in need.⁴ Put differently, the model identifies the impact of resources using the variation in resources not generated by variation in need, which may be capricious or, worse, misallocated—say, due to reliance on inefficient teacher transfer rules or hold-harmless provisions. Notice that if resources are strictly allocated according to formula (as weighted student funding advocates propose), there may be no variation left and the coefficient will not be estimable.

At the same time, one might argue that the remaining variation is essentially random and allows us to estimate the impact of random increases in resources to schools—that this is just the right thing. While this has some appeal, it also means that the estimated coefficient does not provide a meaningful answer to the key policy question: does strategically increasing resources to schools improve student performance? This is the question that matters to policy makers.

Answering this question will require investigating and understanding the political economy of education as well as policy adoption and implementation. What caused the “natural experiment”? What determines eligibility cutoffs used in regression discontinuity designs? Were they set at the point where marginal benefits are expected to be zero (ensuring that no student that might benefit from the program is denied services) or where marginal benefits were thought to exceed marginal costs (consistent with a different cost-benefit analysis)? What explains the processes used to allocate resources across schools within districts (or across districts within states)?

Finally, many important questions do not, in fact, involve causal relationships at all. Some are really about “just the facts.” How big is the black-white test score gap? Is it bigger for boys or girls (or the same)? Has it increased or decreased with NCLB? How many children change schools before the end of the school year or between school years? Is mobility greater among poor students? Is the current wave of foreclosures affecting all groups of students, or are they concentrated in particular schools or districts that may be in need of help? Does school funding increase with increases in students with special needs (i.e., LEP, poor, learning disabled)? How much? These kinds of measures are

4. Alternatively, if additional teachers are routinely allocated to schools based upon the number of students, weighted, say, by the percentage of limited English proficient or special education students, the regression that controls for these is identifying the impact conditional on the variables used to allocate the resources.

critical to good policy making, both because they suggest directions for further research (what explains observed disparities?) and because they indicate how and where interventions might be targeted and most effective (e.g., who needs help?).

BEYOND THE CLASSROOM, THE SCHOOL DAY, AND THE SCHOOL YEAR

Although classroom activities are important, student performance is clearly shaped by a wide range of nonclassroom or nonschool factors, such as housing, health, or nutrition, and out-of-school activities such as after-school programs, summer camp, libraries, or museums. While few would argue that academic outcomes depend only on what happens in the classroom, during the school day, and during the current academic year, much of our empirical research proceeds as if it were so.

To some extent, this reflects limitations in data. Education data sets are understandably thin on variables capturing the noneducational or out-of-school-time features of students' lives. Similarly, data sets assembled for health research or housing research are typically thin on data on schools and education. To some extent, however, the relative scarcity of research in this area reflects the sectoral compartmentalization of government agencies, foundations, and researchers. Where does research on the impact of housing on education (or nutrition and schools) "fit"? Who funds it? Who publishes it?

My own research in New York City suggests that pushing these boundaries is both possible and important. As an example, in a recent study Schwartz et al. (2009) examine the academic performance of students living in public housing in New York City. Assembling the necessary data was time consuming and difficult and required matching data on student residences to public housing projects. The results, however, are intriguing, indicating that students living in public housing are uniquely disadvantaged. To be specific, we find that while there are relatively small differences in the resources provided to schools serving public housing residents, there are significant differences in the student body. More important, children in public housing score lower on standardized tests than otherwise similar children in their schools. This descriptive study raises a host of important questions. Are the observed lower scores the result of conditions in public housing or unobserved characteristics of families living in public housing? Alternatively, are they because of differences in the opportunities and resources in the neighborhoods surrounding public housing? What matters? After-school programs? Libraries? Health clinics? Community centers?

Clearly, health and nutrition matter to student performance. Students with poorer health are more likely to miss school, affecting their academic

performance. Asthma or obesity may affect social and physical well-being, with spillover effects on educational achievement. How does the school food program fit in? While the percent of students eligible for free lunch is routinely used as a measure of poverty in education research, there is little research that focuses on the program directly. Does the price matter? Does the menu? The limited research available suggests that there is much to be learned here.⁵

One of the implications of excluding out-of-school-time activities and outside-the-school resources is that doing so hinders our ability to isolate the impact of school factors on student outcomes. If these activities shape school outcomes, production functions are misspecified. If the quality or quantity of these activities is related to school resources or practices, this misspecification will lead to biased estimates of the impacts of these resources or practices. Returning to the class size example, parents may spend more (or less) time teaching their children depending on the success of the teaching enterprise at school. They may hire more (or fewer) tutors, have their children participate in more (or fewer) after-school programs, or invest in more (or fewer) computer instructional programs. If we find no effect of increasing class size in high school, we will want to make sure to distinguish between the partial effect (which holds other things, such as out-of-school tutoring or test preparation, constant) and the full effect. Providing good policy guidance may require knowing both, but it certainly requires knowing which one you have estimated and the difference between the two.

THE IMPORTANCE OF URBAN SCHOOLS

Although American education needs to be improved across the board, large urban areas deserve particular attention. The first reason is their size. The sixty-seven large urban school districts of the Council of the Great City Schools (CGCS) educate 7.1 million children and employ almost half a million teachers in over 11,700 schools (CGCS 2009). This represents 15 percent of public school students nationally, 14 percent of teachers nationally, and 12 percent of public schools. According to the National Center for Education Statistics (NCES), in 2005–6 the one hundred largest public school districts educated 11.3 million children (22.7 percent of all public school students) in 16,584 schools (16.7 percent of all public schools) employing over 20 percent of all teachers in the United States (Garofano and Sable 2008).

Second, students in urban school districts are more likely to be poor, minority, foreign-born, and limited English proficient than students elsewhere. More than one-third of the students attending CGCS districts are African

5. See, for example, Figlio and Winicki 2005 or Hinrichs 2009.

American (36 percent) and Hispanic (35 percent), and nearly two-thirds are eligible for free or reduced price lunch (61 percent). Further, 17 percent are English language learners (ELLs), and 13 percent have individualized education programs (IEPs). Importantly, these students represent 32 percent of all African American students, 26 percent of all Hispanic students, 23 percent of all poor students, 29 percent of all ELLs, and 14 percent of all students with IEPs.

Finally, students in city schools have significantly worse educational outcomes: their National Assessment of Educational Progress scores are lower in reading and math in both fourth and eighth grades. The NCES reports graduation rates for these districts well below the national average. As an example, in 2005–6, Los Angeles, Chicago, Dade County (Miami), and New York City showed freshman graduation rates below 60 percent.

Thus the inadequate performance of American students and the disparities between blacks and whites or between the poor and the nonpoor (among others) significantly reflect the inadequacies of urban education. Improving education in the cities is a lynchpin to reducing the national race gap in academic achievement and improving educational opportunities for the disadvantaged.

Unfortunately, the research base on urban education is insufficient to guide urban education policy. Although many of the issues facing urban schools and students are universal, some are unique, due in part to the difference in scale. The policies and practices that might be successful or effective in districts with a handful of, or even several, schools may be ineffective in large districts, numbering in the hundreds of schools. As an example, small high schools might be appealing in districts with a few thousand students but prohibitive or unwieldy in large ones; it is impractical to educate 100,000 high school students in 2,000 high schools of 500 students.

Large urban districts are often quite diverse. While Tiebout sorting may work well to deliver relatively homogenous suburbs, cities are diverse. Understanding the performance of urban students means looking not only at blacks and whites but also at Hispanics and Asians, at the differences between native-born and foreign-born students, and at differences within the immigrant community—between students from different countries or those who speak different languages at home.

This challenge also presents opportunities. The size and diversity of urban school districts offer unique variation that we can exploit to explore the causes and consequences of differences in academic performance and an opportunity to understand how and why some schools succeed while others do not. Are schools that serve white students well also good for black students? How about Hispanic students? Further, cities offer the possibility of assembling large longitudinal data sets from administrative data on students, staffing, and

school characteristics, as well as the important related factors—neighborhoods, housing, health, etc.

Notice that understanding urban education will require expanding our theoretical tool kit. Models from public finance explain the behavior of school districts. Labor economics provides models of individual behavior (for teachers, students, principals, parents). The theory of school behavior and the intradistrict decisions of districts and schools, which might draw on industrial organization, is, in contrast, quite thin. How should (do) districts organize schools within a large district? Is there an optimal portfolio of schools or an optimal distribution of resources and/or students across schools?⁶ What are the implications of alternative policies for neighborhoods or the city as a whole?

My own work, with Leanna Stiefel and colleagues, on education in New York City provides an illustration of both what is possible and what needs to be done.⁷ The largest school district in the country, New York City, educates over 1.1 million students in more than 1,400 schools and is incredibly diverse. Over a third of New York’s students are black, a third are Hispanic, and there is a large immigrant population, including students from a wide range of countries and speaking many different languages at home. This diversity is also seen in the wide array of neighborhoods, ethnic enclaves, and distinct communities spread across the five boroughs.

At the same time, there is broad variation across schools. While some are quite small (100–200 students), others are quite large (4,000+ students). Some are virtually all poor and some have hardly any poor students. Some are nearly all black or Hispanic, others are predominantly white or Asian, and still others are fairly integrated. New York City has both award-winning excellent schools and failing schools identified by state and federal accountability programs as “in need of improvement.” Some schools have a long history, and there are also new schools opening each year. More generally, New York’s “experimentation” with policy changes and reforms yields ample opportunities for researchers to learn about what works and what does not.

Equally important, we have assembled an extraordinary data set on education in New York City. This includes more than thirteen years of longitudinal data on more than a million students, including information on testing, sociodemographics, schools attended, and home language, complemented by

6. Rubenstein et al. (2007) begin to examine this by exploring how resources are allocated across individual schools in large districts.

7. I have been fortunate to have terrific colleagues and partners in the New York City research. Although I cannot list all of our New York City projects and authors here, I want to acknowledge the important work of Hella Bel Hadj Amor, Vicki Been, Luis Chalico, Colin Chellman, Dylan Conger, Sean Corcoran, Ingrid Ellen, Patrice Iatarola, Brian McCabe, Charles Parekh, Ross Rubenstein, Ioan Voicu, Meryle Weinstein, Matt Wiswall, and Jeff Zabel.

data on schools, housing, property values, neighborhoods, teachers, and more. Data on elementary and secondary education in New York City public schools are linked to data on the City University of New York, which educates many graduates of the public schools.⁸

In a 2007 study, Leanna Stiefel, Ingrid Ellen, and I examined the distribution of the black-white test score gap across elementary and middle schools (Stiefel, Schwartz, and Ellen 2007). The results were interesting. First, segregation meant that many schools had too few students to have a meaningful “gap.” Among those with significant populations of both groups, we found considerable variation in the magnitude of the gap across schools. While a range of student characteristics explained some of the gap, significant disparities remained. Further, the black-white differential estimated using a school fixed effects model was little changed by using a classroom fixed effects model, suggesting that disparities are not simply due to within-school sorting and inequities across classrooms. Like the research on public housing, this article points to more questions about the underlying causes of these disparities. Are differences due to differences in student health, neighborhoods, or treatment within shared classrooms?

Interestingly, our companion investigation of the educational outcomes of immigrant students found that the “nativity gap” favors immigrants. Schwartz and Stiefel (2006) found that immigrants outperform the native born on state tests and have higher attendance rates and higher graduation rates relative to otherwise similar native-born students (that is, controlling for English proficiency, poverty, and so on). Yes, limited English proficiency means lower test scores, but this is true for both native- and foreign-born students. In current work examining the impact of age of entry on the high school performance of immigrants, we find evidence that mobility matters to the native born as well as to the foreign born (see Conger, Schwartz, and Stiefel 2008). Does mobility explain performance disparities?

Understanding mobility is challenging, in part, because there are many different kinds of mobility: across schools, districts, states, or countries; within and between academic years; school mandated or discretionary; voluntary or involuntary; and so forth. Further, longitudinal data on mobile populations are hard to find. Using data on New York City public school students, Schwartz, Stiefel, and Chalico (2007) found that eighth-grade performance declines with the number of schools a student has attended. At the same time, mobility is highest among at-risk populations, including poor, black, and Hispanic students.

8. See Schwartz and Stiefel 2006 and Stiefel, Schwartz, and Ellen 2007.

Even more, Rubenstein et al. (2009) found that K–8 schools deliver higher test score results than other types of schools, reflecting, in part, the fewer school moves made by their students. While this limited evidence hardly warrants large-scale reconfiguration of elementary and middle schools into K–8s, further exploration of the potential benefits of K–8s and other interventions to limit mobility are clearly warranted.

FINAL THOUGHTS

While this essay has focused on research and policies aimed at improving test scores and academic performance, it is clear that this focus is too narrow. Indeed, the public demand for public education is not confined to a demand for high test scores or graduation rates. Parents and students (and taxpayers and employers) also care about athletics, social and emotional development, civic engagement, arts and cultural education—that is, football games, proms, and band. Ignoring these hampers our ability to identify policies that will garner the political and social support needed for success.

At the same time, we must take care in communicating research results to the policy-making community. Too often heroic leaps are made from narrow results to policy guidance. This ultimately undermines our credibility. It is perhaps no one's fault but our own that we hear policy makers and leaders advocate interventions that we “know work” such as charters or merit pay—even when the research base is thin and conflicting.

In the end, making research in education finance and policy matter means asking the important questions and designing studies that provide thoughtful, nuanced answers that inform policy. As the mathematician John Tukey observed: “Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise” (1962, p. 13).

REFERENCES

- Angrist, Joshua D., and Victor Chaim Lavy. 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics* 114 (2): 533–75.
- Conger, Dylan, Amy Ellen Schwartz, and Leanna Stiefel. 2008. Do immigrants differ from migrants? Disentangling the impact of mobility on high school completion and performance. New York University, Institute for Education and Social Policy Working Paper No. 08-01.
- Council of the Great City Schools (CGCS). 2009. *Urban school statistics*. Available www.cgcs.org/about/statistics.aspx. Accessed 14 August 2009.
- Figlio, David, and Joshua Winicki. 2005. Food for thought: The effects of school accountability plans on school nutrition. *Journal of Public Economics* 89: 381–94.

Garofano, Anthony, and Jennifer Sable. 2008. *Characteristics of the 100 largest public elementary and secondary school districts in the United States: 2005–06*. NCES 2008-339. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Glod, Maria. 2008. Searching for science to guide good teaching. *Washington Post*, 28 April.

Hanushek, Eric. 1986. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature* 49 (3): 1141–77.

Hanushek, Eric. 1999. Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis* 21: 143–63.

Hinrichs, Peter. 2009. The effects of the national school lunch program on education and health. Unpublished manuscript, Massachusetts Institute of Technology.

Ladd, Helen F., and Edward B. Fiske. 2007. *Handbook of research in education finance and policy*. New York: Routledge.

Nechyba, Thomas. 2000. Mobility, targeting and private school vouchers. *American Economic Review* 90 (1): 130–46.

Nechyba, Thomas. 2003a. Public school finance and urban school policy: General versus partial equilibrium analysis. *Brookings-Wharton Papers on Urban Affairs* 139–83.

Nechyba, Thomas. 2003b. What can be (and what has been) learned from general equilibrium simulation models of school finance. *National Tax Journal* 56 (2): 387–414.

Nye, Barbara, Larry Hedges, and Spyros Konstantopoulos. 1999. The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis* 21: 127–42.

Rubenstein, Ross, Amy Ellen Schwartz, Leanna Stiefel, and Hella Bel Hadj Amor. 2007. From districts to schools: The distribution of resources across schools in big city school districts. *Economics of Education Review* 26 (5): 532–45.

Rubenstein, Ross, Amy Ellen Schwartz, Leanna Stiefel, and Jeffrey Zabel. 2009. Spending, size and grade span in K–8 schools. *Education Finance and Policy* 4: 60–88.

Schwartz, Amy Ellen, Brian McCabe, Ingrid Ellen, and Colin Chellman. 2009. Do public schools disadvantage students living in public housing? New York University, Institute for Education and Social Policy Working Paper No. 09-08.

Schwartz, Amy Ellen, and Leanna Stiefel. 2006. Is there a nativity gap? New evidence on the academic performance of immigrant students. *Education Finance and Policy* 1: 17–49.

Schwartz, Amy Ellen, Leanna Stiefel, and Luis Chalico. 2007. The multiple dimensions of student mobility and the implications for academic performance: Evidence from New York City elementary and middle school students. A condition report for the New York Education Finance Research Consortium. Available www.albany.edu/edfin/documents/schwartzstiefelmobilitypaper.pdf. Accessed 14 September 2009.

Stiefel, Leanna, Amy Ellen Schwartz, and Ingrid Ellen. 2007. Disentangling the racial test score gap: Probing the evidence in a large urban school district. *Journal of Policy Analysis and Management* 26: 7–30.

Tukey, John W. 1962. The future of data analysis. *Annals of Mathematical Statistics* 33 (1): 1–67.

Viadero, Debra. 2009. “No effects” studies raising eyebrows. *Education Week* 28 (27): 1, 14–15.