

Testimony of Grover J. Whitehurst

Exploring Opportunities to Strengthen Education Research While Protecting Student Privacy

U.S. House of Representatives

Committee on Education and the Workforce

Subcommittee on Early Childhood, Elementary, and Secondary Education

June 28, 2017

Mr. Chairman and Members of the Subcommittee:

As someone who has been an active researcher and user of privacy protected student data, as well as the founding director of the Institute of Education Sciences, I am very appreciative of the opportunity to testify today.

The mission of IES is to report statistics on the condition of education in the United States; to fund research on educational programs and practices that support student learning; and to evaluate the effectiveness of Federal and other education programs.

The federal education research enterprise, which was established by the Cooperative Research Act of 1954, was a failure prior to the establishment of IES in 2002. Books have been written about that. I will summarize with a quote from a National Academies of Science report on the immediate predecessor to IES, the Office of Educational Research and Improvement. After cataloging a litany of problems with the agency, the NAS committee concluded that:

“OERI needs to be rebuilt.”ⁱ

Congress did that with the Education Sciences Reform Act of 2002. Here is what the Office of Management and Budget said in 2008 in its program assessment rating of IES:

Since its creation by the Education Sciences Reform Act of 2002, IES has transformed the quality and rigor of education research within the Department of Education and increased the demand for scientifically based evidence of effectiveness in the education field as a whole.ⁱⁱ

This leads me to the first of four points and recommendations:

1. **If it ain't broke, don't fix it.** The operational success of IES in the 15 years since it was founded has spanned two long-serving directors, several leaders delegated the responsibilities of the director, and three different presidential administrations. This stability in performance across variations in leadership is due to design features in the Education Sciences Reform Act. These include a clear and focused mission; statutory independence from political interference in

research and reports; highly qualified staff; strong internal controls for quality of products; and predictable funding.

I recommend that as you take up the reauthorization of IES, you retain the core components of the Education Sciences Reform Act of 2002, and focus on ways to add functions to IES that will make it even more useful.

2. **Follow the money.** IES operates with an annual budget of a little more than \$600 million. This amount is small in relative terms -- only about 1% of the Department of Education's discretionary budget. But on the bright side, the IES budget has risen over the years and has not whipsawed. My immediate concern is not so much with the line item program budget of IES, but with its budget for salaries and expenses, which is drawn from the Department's omnibus budget for administration and is at the discretion of the Secretary. IES has substantial and absolutely essential independent statutory authority to carry out its work, but it would be easy for an administration displeased because IES was reporting findings that were off-message, or for an administration that simply had other priorities for its administrative budget, to cripple IES through reductions in funding for its salaries and expenses. This danger is evident in the Department's present plans for an across-the-board reduction in force that may cut IES staff, already small in number and operating with high efficiency, to the bone.

I recommend that in reauthorizing IES, this Committee provide for a specific budget for administration of IES. The funding for this would come from a corresponding reduction in the Department's omnibus budget for administration. This would strengthen the independence of IES from political control compared to the present arrangement in which the allocation of funds to cover the staffing of IES depends on the good will and priorities of the administration in power.

3. **Time is of the essence.** The yield in terms of usable products from individual IES investments typically takes years to emerge, e.g., a research firm gets a multiyear contract to evaluate the impact of a big budget federal program and a public report of that work is released four years later. There is a natural timeline of these activities that can't be cut short, and the work is critically important. At the same time, a lot of education policy making occurs in time frames measured in months rather than years and, presently, is usually bereft of high quality, objective research evidence that is summarized in ways that will inform decision makers. Consider, for example, the Trump administration's FY18 budget proposal to cut funding to colleges and universities for the federal work-study program. This Congress will need to consider that proposal. Wouldn't it be great if IES were funding a program activity that, in a matter of months, could generate an objective report on what we know and don't know about how federal work study is performing, who it serves, and what its consequences are for participating students?

I recommend that you authorize a budget for IES to generate independent quick turn-around reports on high impact policy issues that arise in education policy at the federal and state levels. This could be paid for by repositioning funds elsewhere in the Department's administration or national activities budgets that are typically spent on related but non-politically independent activities.ⁱⁱⁱ

- 4. Invaluable information to improve education is found in large administrative databases of student records. Mining it has risks that must be acknowledged and eliminated.** The computational analysis of large administrative data sets can reveal patterns, trends, and associations that provide important insights on how to improve education policy and practice.

There are two things that stand in the way of the wide utilization of these databases to fuel smart education reform. The first is legitimate concerns about the protection of the privacy of individual student education records. No parent wants their child's identity and test scores in 6th grade available to hackers, or released on the internet by mistake, or expropriated for commercial purposes by businesses with which school districts enter into contractual relationships. The protection of the privacy of digital records, whether educational or otherwise, is not child's play. I've been told on the Q.T. by state education officials that they worry greatly about the security of their statewide longitudinal student records because they don't have in their offices and can't afford to hire from outside the expertise to secure them. This has to be fixed before this Congress or the public you represent should be comfortable with the merging of student records to support large scale data analytics.

The second thing that has to be fixed, which is bound to the solution of the privacy issue, is the availability of student records to education researchers and policy analysts outside of state education bureaucracies. Through the Statewide Longitudinal Database Grant program, funded by Congress and carried out by IES, the states have received \$721 million to support their student data systems.^{iv} The legislative authority for these grants, found in Title II of the Education Sciences Reform Act, calls for the funds to be awarded to support data systems "to facilitate research." Forty-seven states have received awards under this program. Using a liberal standard, maybe three of them allow ready access to facilitate research. Every other state has the data locked away or available with heavily burdensome obstacles. States are, perhaps understandably, not eager to give independent researchers access to data that may lead to findings that call state policies and practices into question. And they are aware of their data security issues and their lack of expertise in how to share data so that it is truly de-identified, e.g., so that Susie's test scores are in the database but no-one can link them to the person of Susie. It is just easiest and safest for states to say no to requests for access.

I recommend that in reauthorizing IES and in budgeting for another round of funding of the Statewide Longitudinal Data Systems grants this Committee support IES in carrying out a competition in which funds are available to states to establish data centers that have frontline technical responsibilities for insuring the privacy of individually identifiable student data from

external threats and for enabling easy access by outside researchers to non-personally identifiable longitudinal data (de-identified and anonymous).^v The North Carolina Education Data Center at Duke University provides a longstanding example of how states could accomplish this.

Thank you for the opportunity to testify. I look forward to your questions and comments.

ⁱ <https://www.nap.edu/read/1973/chapter/2#3>

ⁱⁱ <https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/expectmore///rating/perform.html>

ⁱⁱⁱ For instance, the [Office of Planning, Evaluation, and Policy Development](#) carries out evaluation and planning studies to support the Secretary's priorities. Frequently these studies overlap functions that are fulfilled by IES, which, unlike OPEPD, is independent of political control and has high standards for quality and objectivity. Funds that flow through OPEPD for IES-like activities would be more productively spent by IES on quick-turnaround public reports summarizing objectively what is known from research relevant to impending education policy decision making.

^{iv} https://nces.ed.gov/programs/slds/ppt/SLDS_PPT.pptx

^v There are several ways to provide non-personally identifiable data to independent researchers that go far beyond simply stripping obviously personally identifiable information out of a database of student records. For example, researchers can be required to analyze data through a web program that can only produce aggregate results and that does not involve any transfer of or access to individual data. Thus, the researcher can run a regression analysis on data of interest without ever seeing or having access to the data on which the analysis is run, which, itself, has already been de-identified before being provided as the backend to the web-based analysis tool. The general public as well as researchers can do this today for NAEP data on the IES/NCES website with the [Data Explorer](#).