

**Institute for Policy Research  
Northwestern University**

**Grover J. "Russ" Whitehurst**

Director  
Institute of Education Sciences  
U.S. Department of Education

**IPR Distinguished Public Policy  
Lecture Series  
2003-04**

**"Making Education Evidence-Based:  
Premises, Principles, Pragmatics,  
and Politics"**

**April 26, 2004**

Editor: Patricia Reese  
Copy editor: Audrey Chambers  
Editorial assistant: Meredith Buse  
Layout: Alice Murray. Original design: Valerie Lorimer  
Photographs: Jean Clough

© 2004 Northwestern University. All rights reserved.

---

## Foreword

Fay Lomax Cook, Director

Grover J. “Russ” Whitehurst, IPR’s 2004 Distinguished Public Policy Lecturer, helped to found the Institute of Education Sciences within the Department of Education in order to transform education into an *evidence-based* field. As he explains his initiative in these pages, the guiding premise is that the key to progress in education is scientific research and evaluation together with systematically collected data on education performance.



This aim perfectly coincides with IPR’s mission—to bring excellent social science research to bear on important social problems and on social policy decision making. As Whitehurst points out, evidence-based policymaking is already established in a number of fields in the U.S. such as health care and agriculture. But it does not have a strong place yet in education, and that is what he is trying to change. He aims for a time when decision makers “routinely seek out the best available research and data before adopting programs and practices that will affect significant numbers of students.”

In his lecture, Whitehurst describes why progress in education requires scientific research. He then discusses the principles that underlie evidence-based research in education: Progress requires scientific research; education isn’t unique; methods matter; and usefulness is paramount. Finally, he examines the pragmatics and politics of transforming education into an evidence-based field.

Whitehurst is at the center of an effort to transform the way education research is conducted. This lecture describes the promises and the challenges of that effort. In his attempts to bring research about what works to bear on policy discourse and decision making, Whitehurst is an exciting example of those who successfully bridge the gap between the worlds of social science and policymaking.



## Grover J. “Russ” Whitehurst



Grover J. “Russ” Whitehurst was appointed by President George W. Bush to a six-year term as the first director of the Institute of Education Sciences (IES). The institute was established within the U.S. Department of Education by the Education Sciences Reform Act of 2002. The IES conducts, supports, and disseminates research on education practices that improve academic achievement, statistics on the condition of education, and evaluations of the effectiveness of federal and other education programs.

As director, Whitehurst administers the institute, including the activities of the National Center for Education Statistics, the National Center for Education Evaluation and Regional Assistance, and the National Center for Education Research. He coordinates IES’s work with related activities carried out by other agencies within the department and the federal government. He advises the Secretary on relevant research, evaluation, and statistics. And he engages in many activities to encourage the use of scientifically based research in education policy and decision making throughout the U.S.

Whitehurst earlier served as assistant secretary for the Office of Educational Research and Improvement, the institute’s predecessor. In that role he established the What Works Clearinghouse, initiated new programs of research such as those in reading comprehension and preschool curriculum, upgraded the rigor of scientific peer review, promoted the use of scientific evidence throughout the Department of Education, and spearheaded a historically unprecedented increase in the presidential budget request for education research.

Previously, he was Leading Professor of Psychology and Pediatrics and Chairman of the Department of Psychology at the State University of New York at Stony Brook. Whitehurst has authored or edited five books and published more than 100 scholarly papers on language and prereading development in children.

**“Making Education Evidence-Based:  
Premises, Principles, Pragmatics, and Politics”**

**Grover J. Whitehurst, Director  
Institute of Education Sciences  
U.S. Department of Education**

I appreciate very much your invitation to deliver the 2004 Distinguished Public Policy Lecture. It has provided me with an opportunity to step back from the day-to-day responsibilities of my job and think more generally about where the nation finds itself with respect to education research and policy and the challenges ahead.

This is a particularly inviting venue because my thinking about the nature and role of evidence in public policy has been heavily influenced by the work of the so-called Northwestern school of evaluation. Cook and Campbell’s (1979) Quasi-experimentation was a staple of the graduate research methods course I taught for years at the State University of New York at Stony Brook. And faculty and alumni of the Institute for Policy Research have been frequently involved in the various technical working groups, advisory committees, and peer-review panels that help us plan and carry out the work of the Institute of Education Sciences (IES). So if there is anything about the direction of IES or what I have to say today that you don’t like, you have no one but yourselves to blame for it.



We are at the beginning of the transformation of education into an evidence-based field. By evidence-based, I mean an endeavor in which decision makers routinely seek out the best available research and data before adopting programs or practices that will affect significant numbers of students.

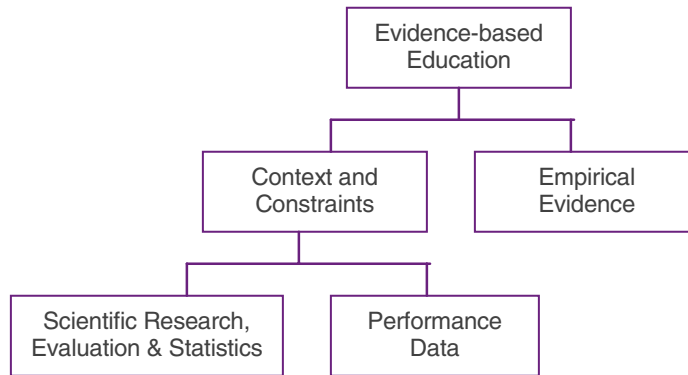
My role at the U.S. Department of Education for the last three years, with the help and support of many inside and outside of government, has been to move the field of education towards a tipping point after which current operating modes will be replaced with empirical ones. That has required commitment to a clear set of principles, pragmatic action to advance them, and sensitivity to the formal and everyday political context in which that action is embedded.

### Premises and Principles


Let me describe some of those principles and premises, with the push-back they generate, as a way of highlighting what IES is trying to accomplish and the challenges of doing so.

#### *1. Progress requires scientific research*

Our guiding premise is that scientific research and evaluation, linked with systematically collected and utilized data on education performance, is the key to progress in education. Indeed, we assume that evidence-based practice and policy is the best—and perhaps the only—way to produce continuous improvement in education outcomes.



This figure provides a schematic of the distinctions between scientific research and performance data, and how we see them fitting together in evidence-based education. In brief, scientific research, evaluation, and statistics are produced by scientists and typically




appear in peer-reviewed journals and other outlets that are read by a technical audience. For instance, a scientific evaluation might indicate that a particular reading practice is effective in increasing children's decoding skills when compared with business as usual in a randomized trial. Performance data, in contrast to scientific research, is produced by school systems and other entities that deliver education to determine whether the programs and practices that have been deployed are meeting goals. A school that adopted the reading practice that had been demonstrated to be efficacious in a scientific evaluation should collect data on how children are performing in the classrooms using that practice to identify whether the program is working as deployed and to address potential problems, such as weak implementation. Together, scientific research and performance data comprise empirical evidence.

In evidence-based education, decision makers consider empirical evidence in the context of practical constraints: Imagine that reading program A has been shown in several well-designed studies to be somewhat more effective than reading program B. However, the costs of adopting and supporting A are higher than those for B. With both costs and effectiveness in mind, a district chooses B rather than A as its new reading program. That is a respectable evidence-based decision.

Opposition to the premise of evidence-based education, as outlined here, is widespread. It is not confined to those who are relatively uninformed about scientific research and evidence-based decision making. There are many critics within the academy:

In the past—indeed, in the present—much of the best school practice has been based on ... seat-of-the-pants observations, reflections, and informal experimentation. Perhaps we need to be doing more of this, rather than less; perhaps, in fact, research dollars might be better spent on setting up teacher study groups or mini-sabbaticals, rather than on NIH-style field-initiated or targeted-grant competitions. (Gardner, 2002)



I've described evidence-based education as a premise because we're going to proceed on that basis. There is no need to conduct an experiment in which we, for example, set up a second federal education research agency, give it half our money to support mini-sabbaticals and seat-of-the-pants observations, and wait 10-15 years to find out whether that approach generates more progress than systematic empiricism. Such approaches have resulted in little progress. We're committed to systematic empiricism as the basis for evidence-based education, and that is the path we will follow.

## *2. Education isn't unique*

We operate on a premise that the relation between education research and practice is similar to the relation between research and practice in other fields that involve human behavior. This doesn't mean we believe that education is just like clinical psychology or health care or social welfare or violence prevention. Rather, we assume there is enough overlap between these fields and education that we can profit from their methods and approaches and we can learn from the history of their transformation into evidence-based endeavors.

As with our premise that progress requires research, there is considerable opposition to the view that there are parallels between education and fields such as health care. Here is one example from a senior state education official:

This emphasis on a medical model for education research is abhorrent.... Our children are not sick or diseased. Education and instruction are not treatments. (Viadero, 2002)

The "medical model" referred to by this critic is the use of rigorous experimental methods, such as the randomized clinical trial, to determine what works best for whom under what circumstances. Although experimental methods have, in the last 50 years, come to dominate the determination of the effectiveness of drugs and medical procedures, randomized experiments are not an invention of medical science. As most in the audience know, the origins of randomized





trials are in agriculture. The crossover into medicine came rather late, and there are thousands upon thousands of randomized trials on human behavior in fields such as psychology and social welfare.

Even in health care, there are very large numbers of randomized trials that focus on complex human behavior as outcomes rather than disease. For instance, the Cochrane Collaboration, which summarizes results of effectiveness trials in medicine, lists 72 reviews involving education as the intervention—that is 72 reviews, not 72 studies. Each review covers multiple studies.

For instance, a Cochrane review of the literature on education interventions for schizophrenia identified 10 studies in which the treatment consisted of efforts to increase the knowledge and insight of patients with schizophrenia and their family members into the nature of the disease and its treatment. Evidence from these trials indicated that educational approaches had positive effects on both compliance with treatment and symptoms.

There is, in short, nothing about the logic or application of randomized experiments that requires the assumption that the outcomes under investigation are symptoms of disease and that the interventions being studied are pharmaceutical. The randomized trial is as relevant to education as to medicine, and there are thousands of examples of its use in studies in which the intervention is social and the outcome is behavioral.

Another assertion by those who argue that education is unique is that every child in a classroom is different from every other child. We can grant that assumption, just as we can grant that every medical patient is different from every other medical patient, without jumping to the false conclusion that a common intervention won't have a net effect across this variation. Instead of thinking of the well-trained teacher, or primary-care physician, as an artist whose professional actions are creative expressions, the evidence-based perspective defines their role as implementing and monitoring the success of research-validated protocols, and making adjustments as necessary to achieve the best outcome for the individuals under their care.

Another assumption made by those who think education is unique is that schools are very complex institutions that do not support




the uniform implementation of programs and practices. As this argument goes, the complexity of schools leads to weak program implementation. Weakly implemented programs can't overcome the strong effects of students' home and genetic backgrounds. Thus the effects of schools and schooling founded on evidence-based practice



will necessarily be too small to generate substantial gains in education outcomes.

This point is unpersuasive because it flows from false assumptions about other fields. The proposition that the net effect of most education interventions is likely to be relatively weak is probably correct. However, the effect sizes of medical interventions, e.g., hormone replacement therapy, are typically substantially smaller than those that, by convention, are characterized as “small” in social-behavioral-educational interventions. Fields in which outcomes have multiple determinants typically progress by understanding and controlling many relatively small effects. It is the systematic attention to each of these influences and their combination into intervention delivery systems that cumulates in medically or educationally significant improvements in patients' or students' lives. An unfortunate and unavoidable consequence of the view that education progress flows from seat-of-the-pants observation rather than systematic empiricism is that the ability to detect relatively small effects and to examine how intervention effects are influenced by characteristics of students, teachers, and settings is lost or substantially diminished.

Those who hold that education is unique and that scientific research will not transform it are akin to 19<sup>th</sup>-century Luddites. These activists held uprisings against the advances of technology in textiles and agriculture that threatened their way of life. The technologies of systematic empiricism, including the randomized trial, statistical modeling, psychometric assessment, and quantified observations are a threat to the way of life of researchers who are not trained in these technologies, and to education professionals whose practices are




grounded in pre-empirical professional wisdom, intuition, and self-directed creative expression. It is not surprising that they oppose a direction that threatens the status quo.

Who is right? Those committed to a view of education as a unique art and craft, or those committed to education as an evidence-based enterprise? We can't know for sure, but education, a field still largely prescientific, has shown little improvement in productivity and progress in the last half century. The picture is very different in fields that have turned from professional wisdom to systematically gathered and analyzed evidence—for example, agriculture, health care, and clinical psychiatry/psychology.

### *3. Methods matter*

Implicit in the premise that progress requires research and that education isn't unique is the assumption that methods matter. The history of other fields that have become grounded in science shows a progression from decision making based on eminence, to decision making based on evidence derived from systematic protocols for collecting and analyzing data. In medicine, for example, randomized trials to support claims of clinical effectiveness were first required by the federal Food and Drug Administration as a condition for the introduction of new drugs into the market in the mid-60s. Prior to that, the FDA's role had been to prevent the entry of unsafe drugs into the market, with heavy reliance on the opinions of leaders in the field, i.e., eminence. However, it was challenging at best to translate into government action the convictions of eminent clinicians who were frequently in disagreement. In the absence of standards for judging evidence, professional consensus was elusive, and in the absence of consensus, action was impossible. The FDA's policy commitment to the randomized trial as the arbiter of effectiveness resulted in an explosion of studies using that method, a concomitant period of rapid progress in health care, and a grounding of medical practice in evidence.

That methods matter, and the consequences of the frequent use of weak or inappropriate methods in education, are highlighted in recent newspaper accounts of research on the effects of state accountability systems on student academic achievement.



The first article I will highlight appeared on the front page of *The New York Times* on December 28, 2002. The headline was: “Make-or-Break Exams Grow, But Big Study Doubts Value.” The following excerpt captures the gist of the piece:


Rigorous testing that decides whether students graduate, teachers win bonuses and schools are shuttered, an approach already in place in more than half the nation, does little to improve achievement and may actually worsen academic performance and dropout rates, according to the largest study ever on the issue. (Winter, 2002)

That seems clear enough, but wait—the *Times* spoke again only four months later. The headline this time was “New Ammunition for Backers of Do-or-Die Exams,”

Two new studies make the case that do-or-die exams—which decide whether students graduate, teachers are dismissed, or schools are shut in more than half the states in the nation—have brought about at least a modicum of academic progress, especially for minority students who may get scant attention otherwise. (Winter, 2003)

Each of the studies covered in the two *Times* articles was based on analyses of essentially the same data: the correlation between changes in student scores on the National Assessment of Education Progress and the introduction by states of assessment systems with consequences for students, teachers, schools. How could the same data support such different conclusions? A reporter for another publication, *Education Week*, posed that question to one of the authors of the first study reported on by the *Times*. The researcher answered:

I’ve had a lot of people reanalyze our data ... and each and every one of them have come up with different results. (Viadero, 2003)




I submit to you what is obvious to me: We cannot base education policy on research methods that generate as many answers to the same question as there are researchers addressing that question. Methods matter.

At IES, we identify and support methods appropriate to four categories of questions:

“Descriptive questions” address the state of education—how are students performing, how much are teachers being paid, what do teachers do in the classroom, how do parents feel about the quality of their children’s education, and so forth. Appropriate methods include large-scale assessments, surveys, analyses of institutional records, coded observations of classroom interactions, structured interviews, and case studies, among others.

“Association questions” address the statistical connections between education conditions and education outcomes—how are children’s preschool abilities related to their elementary school achievement, how does postsecondary enrollment change as a function of economic conditions, what is the relationship between family income and parental aspirations for children’s achievement, and so forth. Appropriate methods range from computing simple correlations between two variables to mathematically sophisticated statistical models for simultaneously considering the associations among many conditions and variables.


“What works questions” address the causal effects of programs, practices, and approaches on education outcomes—does a reduction in class size improve student achievement, does systematic phonics instruction in the early grades prevent reading difficulties, what are the effects of teacher-induction programs on teacher practice, and so on. IES has taken a clear and strong position that randomized trials, in which participants are assigned by lottery to experience two or more conditions, are the best method for determining what works. The mean difference between outcomes for participants in conditions being compared in well-conducted randomized trials is an unbiased estimate of the effects of those treatments. Other methods for determining what works provide less certainty that the comparison between conditions is fair and that the resulting estimate of



effects is unbiased. The stronger of the other methods, e.g., quasi-experiments with groups that are well-matched at pretest, are often used when randomized trials are impossible or impractical. But because the causal conclusions from such studies are less certain, it is risky to assume these methods produce answers that complement the results from randomized trials. Some have shown that quasi-experimental methods produce answers that vary widely from randomized trial results, and in unpredictable directions (Glazerman, Levy, and Myers, 2003). Although no single research study is ever definitive, and although randomized trials can be challenging to conduct, they move us toward responsible evidence-based policy more quickly and more efficiently than quasi-experiments and other approximations of randomized trials. Randomized trials remain the best method for producing answers that are the most accurate and trustworthy estimation of impacts. That is why they are the preferred method at IES for addressing what works questions.

“Why questions” address the underlying mechanisms and processes by which causal effects occur. Why do quality preschool programs enhance academic and life outcomes for at-risk students? Is it because children in those programs enter school with higher levels of preparation for academic tasks, or because they develop more positive attitudes about school, or because their parents become more involved in their schooling, etc.? Answering such questions can be practically important because knowing the active ingredients of complex, expensive interventions can open the door to the design of more efficient and effective programs. Methods appropriate to answering why questions include many of the methods exemplified in the previous discussion of descriptive, correlational, and causal questions. Randomized trials, for instance, can be used to test hypotheses about particular causal mechanisms, and sophisticated mathematical models of the relations between multiple variables can be used to identify possible paths of influence between cause and effect.

This brief discussion of four categories of questions is intended, in part, to make it clear that no single method characterizes good science. Rather, it is the degree of match between the method, the question, and the conclusion that is at issue.



Methods matter in all sciences. But as the subject matter becomes more complex, multivariate, and politically grounded, as is the case in education, methods become our only reliable defense against the seduction of desire, faith, hope, and vested interest.

#### *4. Utility is paramount*

IES conducted a survey of a purposive sample of our customers to determine what they think we ought to be doing to serve their needs (Institute of Education Sciences, 2003). The sample included school superintendents and principals, chief state school officers, and legislative policymakers. One question we asked was:

What could the U.S. Department of Education do to make education research more useful, more accessible, or relevant to your work?

Their answers suggest that something needs fixing in the type of work that is conducted by the education research community. For example, 23 of the 30 school superintendents and local education officials who were interviewed spontaneously criticized existing research for its overly theoretical and academic orientation. A typical response was:

There may be less than one percent of the existing research that's really meaningful to teachers. Much is for researchers, for getting funding, for career advancement, or for advocacy. I don't want theories. Teachers need strategies, practices. Give them things that can help teaching and learning, things that can help kids.

We asked these same respondents to identify the issues on which they want good research conducted. They told us they want to know, for example, how to structure a teacher induction program to enhance retention and teacher performance, which of the commercially available mathematics curricula are effective in enhancing student learning, how to design an assessment and accountability system so that negative effects are minimized, how they can structure teacher compensation to attract and retain the best and the brightest, and so forth.




The principle and premise that IES derives from the results of this survey is that education is never going to be transformed into an evidence-based field unless the education research community produces applied research that helps educators solve problems. Methods matter, so the research must be rigorous, but that research also has to be relevant to practice.

The opposition to applied research in education comes from the research community. The most principled opponents argue that in the history of other fields, application has been built on fundamental understanding derived from basic research and use-inspired basic research. Such research, they argue, takes many years and a substantial investment. Without it, they feel, education research is doomed to be a weak affair built on a study of the effects of black boxes.

Basic research in the disciplines related to education, such as economics and psychology, has been ongoing for over a century, has produced basic knowledge relevant to education, and will continue to do so. Some federal research agencies, by statute, are primarily about the business of basic research and the search for fundamental knowledge. The National Science Foundation, for example, has a mission “to promote the progress of science.” IES, in contrast, is primarily about practical action, solving real-world problems, and providing useful information to the public at large.

Without in any way diminishing the value of basic research, our premise is that progress and fundamental understanding can also be generated by research that directly addresses real-world problems. We believe that such research may be particularly needed in areas such as education, in which problem solutions are richly multivariate and contextual. When problem solutions are situated in the changing circumstances of schools, students, teachers, and government policy, even the strongest findings from basic research on learning, instruction, and human interaction will face a complicated and uncertain translation into education practice. Yes, the world needs basic research in disciplines related to education, but education won’t be transformed by applications of research until someone constructs systems and approaches and packages that are engineered to work





in the settings in which they will be deployed. A primary role of IES is to promote such research—research that has high utility to policy-makers and educators.

### **Pragmatics**

Premises and principles define direction, goals, and mission. They are very important but require machinery and action to be accomplished. The pragmatics of transforming education into an evidence-based field has occupied and continues to consume much of my and my staff's time and energy. This is backstage work that I doubt is of much interest to most of this audience, so I will touch only lightly on it.

It has involved changing the structure and nature of the organization, improving the process by which proposals are selected for funding, modifying the focus of our research and evaluation efforts, improving our ability to disseminate research findings, and increasing congressional support for funding education research. All of this has been in service of the goals of increasing the supply and utilization of education research that is both rigorous and relevant to education decision makers.

My initial challenge and pleasure upon arriving in Washington as assistant secretary for research and improvement was to work with Congress as it authorized a new research entity within the Department of Education. On November 5, 2002, the president signed into law the Education Sciences Reform Act. That act replaced the Office of Educational Research and Improvement with the new Institute of Education Sciences. Unlike its predecessor, the IES focuses solely on research, evaluation, statistics, and dissemination. The new legislation provided IES with the flexibility to deploy its resources strategically into areas of greatest need. IES was given the responsibility for the evaluation of the impact of federal education programs, a function previously lodged within the office of the Education Secretary. A new degree of independence was afforded to IES by having the person in my position, the director, and each of three commissioners under the director, serve for six-year terms.



To support these changes, IES was given an “excepted service authority,” which allows us to recruit a significant number of scientific and professional employees outside of the regular civil service system. Using that authority, over 25 top-notch scientists have joined the agency during my tenure. This has allowed us to create a culture of science within the institute that supports high-quality research, evaluation, and statistics.

IES developed more rigorous standards for the quality of funded proposals. We began our efforts to improve the peer review of




research proposals by articulating clear standards in requests for grant applications. Then we implemented new procedures for peer review of applications for research funding that are modeled on those used at the National Institutes of Health. We established a performance tracking system for our research investments by

submitting each year to the same external panel of distinguished scientists a random sample of our new funded grant proposals for an evaluation of quality.

Taking the year before IES was established as the baseline, we have achieved a 94 percent improvement in the rated quality of our funded work over the last two years. Because randomized controlled trials provide the most rigorous tests of what works in education, and because our customers are predominantly interested in questions of what works, we have also tracked annually the proportion of our funded proposals addressing what works questions that use experimental methods. Again, using the year before IES was established as the baseline, the proportion of our funded projects asking causal questions that use randomized experimental designs has increased by more than 200 percent.

To increase the relevance of IES’s research and evaluation activities, we have placed a much greater emphasis on conducting research on the effectiveness of specific programs and practices.




Rather than holding open competitions in which researchers could submit applications for funding for any topic of interest to researchers, IES has established focused competitions in areas in which sustained research was needed and which offered the potential of solutions to major problems in education. Seven new focused research programs have been established and are receiving ongoing funding: preschool curriculum; teacher quality, socialization, and character development; mathematics and science education; school finance, management, and leadership; cognition and student learning; and reading comprehension. Additional focused research programs are planned.

To increase the supply of rigorous and relevant research, we have sought and obtained additional funding from Congress. The 2004 budget for research and dissemination within IES is 60 percent greater than it was in 2000, and the overall budget for research and dissemination, statistics, and assessment is 66 percent greater than it was in 2000. This understates the increased investment, because funds for our evaluation activities come from other program offices in the department and thus are not in the IES's line-item budget.

Recognizing that there are significant capacity issues within the education research community, we have established a program to fund interdisciplinary research training programs in the education sciences. Grants are going to institutions that put together a program across departments such as psychology, political science, economics, education, and epidemiology that provides intensive training in education research and statistics. Predoctoral students will graduate within a traditional discipline, e.g., economics, but will receive a certificate in educational sciences, and will be expected to conduct dissertations on education topics. We have also established new post-doctoral training grants to allow additional opportunities for training and retraining researchers.

Rigorous research by itself will not transform education into an evidence-based field. The knowledge generated by research must be disseminated in a clear, user-friendly, and easily accessible format. To this end, IES created the What Works Clearinghouse. Its sole purpose is to deliver solid research into the hands of educators, policymakers, and the public. To achieve this goal, the clearinghouse




screens and evaluates research studies to identify those that provide trustworthy information on the effectiveness of programs, products, and practices that are intended to enhance student outcomes. The clearinghouse makes this information available through its Web site, <http://whatworks.ed.gov>. The clearinghouse is unlike any previous effort to vet research studies in the social, behavioral, and educational arena in that it depends on transparent standards and protocols, rather than vaguely articulated judgments by particular committees of professionals.

These are only highlights of what has been accomplished to date from an intensive effort to build a research office in the Department of Education that is up to the serious challenge of transforming education into an evidence-based field. There is much work yet to be done, but I believe there is consensus among observers of this process, even those who are critical of some of the directions that have been taken, that IES is different from what preceded it, and that the Department of Education's research office is having an impact on the enterprise of education research and its utilization.

### **Politics**


Members of the research community often assume that any political involvement in research is inappropriate. That perspective loses sight of the source of funding for research, the U.S. taxpayer in the case of IES. And it ignores, in the case of education research, the intensely and appropriately political nature of education itself. Questions of what, when, and how students are to be taught, and with what resources are decided at the local, state, and federal level by elected officials, and by educators who are directly elected or appointed by elected officials. Those political decisions indirectly determine the research priorities for IES to the extent that we intend to conduct applied research to answer questions of relevance to educators. Basic researchers who carry out work that intends to cut at the "joints of nature" have some reason for moral outrage if political action determines topical priorities for funding. But education is not at the joints of nature. It is culturally defined and transmitted. From a research perspective, it is a set of tasks to be solved, with many if



not most of those tasks created in the political arena. So when Congress, for example, passes a piece of education legislation that makes supplemental services available to children on a wide scale, this generates a priority for education research to answer questions about what works best in supplemental services for what children under what circumstances. Congress and state legislatures don't generate the phenomena that biochemists study. They do in education. Education researchers who feel that they—rather than politicians—should determine what is important to study misconstrue their field as a basic science discipline in which scientists are best equipped to know what is important.

Although politics is inextricably bound up with policy and research priorities, it should not have more than a broad oversight role in carrying out and vetting the results of education research. Those are technical tasks best left to those trained to carry them out. Thus Congress and the executive branch are carrying out their legitimate roles in determining that research on the effectiveness of preschool curricula in preparing children for K-12 education is worth funding and at which level. However, they would be intruding inappropriately and detrimentally in dictating a particular research design for such research, or in exercising review and approval authority over publications that would flow from such research.

Congress recognized the critical distinction between political involvement in setting priorities versus political involvement in the conduct and reporting of education research in the Education Sciences Reform Act of 2002. Among the provisions of the statute that intend to guard IES from inappropriate political intrusions are an independent publication authority, which allows IES to publish reports without review or approval by other offices of the Department of Education; a delegation clause, which directs the Secretary of Education to delegate to the IES director all authority necessary to carry out the statute (other than administrative services such as those provided by attorneys, computer technicians, and so forth); and a six-year term for the director and commissioners. The statute also provides for an independent, nonpartisan National Board for Education Sciences that approves priorities proposed by the director after public comment,



can review any grant or contract entered into by IES, and that makes periodic reports to Congress on the functioning of IES.

There is, of course, a world of difference between the formal role of politics in education research as described in statute versus the everyday politics of advancing the agenda of evidence-based education. On the positive side, everyday politics involves being as responsive and useful to elected and appointed political officials as possible, consistent with statutory requirement to avoid partisanship. Thus IES does fact-checking of speeches of senior officials, produces syntheses of education research relevant to ongoing policy deliberations involving the Department of Education, provides technical assistance for congressional committees, works with state-level officials and committees seeking guidance on research evidence, and, in general, treats elected officials and their staff, regardless of political affiliation, as high-priority, extremely valuable customers.

The everyday politics of transforming education into an evidence-based field also involves having a good sense of the political landscape in order to navigate successfully toward the desired destination. This is a partial list of what I perceive to be major features in that landscape:

*Congressional skepticism.* Congress has invested modestly in education research for many years. They find useful their investment in statistical surveys and assessments, such as NAEP and the Schools and Staffing Survey. However, they are skeptical about the value of their historical investment in research and evaluation—viewing much of it as irrelevant or thinly veiled advocacy. The IES commitment to rigor and relevance in our research and evaluation portfolios is, in part, in response to this political reality.

*Policymaker indifference.* Although the rhetoric of scientifically based research and evidence-based policy is popular, most education policy continues to be grounded in intuition and political calculus rather than evidence. Too many policymakers don't bother to check on the evidence before moving ahead, or discount the evidence if it is unattractive to them. There is no overnight solution to this. One useful tactic is for IES to look for and embrace policy




deliberations where minds are not made up, where there is more than one politically viable option, and where evidence can be a useful shield for policy decisions that would otherwise be attacked as self-interested. In such situations, it is appealing to policymakers to be able to say, “We consulted the experts, and they told us that the evidence favored the decision we made.” Of course, if this isn’t to be just another form of advocacy, the research that policymakers use to explain their decisions has to be rigorous and supportive of those decisions.



*Questioning the speaker*

*Policymaker zeal.* The flip side of policymaker indifference is the institution by policymakers of requirements for decisions derived from scientifically based research in advance of that research being available. For instance, a recent review by the National Research Council found that studies of the effectiveness of widely available mathematics curricula, including 19 curricula funded by the National Science Foundation, “fall short of the scientific standards necessary to gauge overall effectiveness.” (National Research Council, 2004). In other words, we do not currently have available rigorous research on what works best in mathematics education. Yet there are laws requiring that mathematics curricula and methods be selected based on scientifically based research. IES’s approach to this problem is to encourage a two-track approach to program and curriculum selection based on evidence. The first is to use that evidence when it exists. The second is to provide incentives for the collection of such evidence when it does not exist.

*Negative results.* Bad news is by definition disappointing to someone. Sociologist Peter Rossi long ago articulated a principle now known as Rossi’s law: The expected value for any measured effect of a social program is zero. To the extent that education programs follow Rossi’s law, the results of rigorous evaluations are




frequently going to disappoint someone. Sometimes the disappointed have trouble separating the message from the messenger.

Our response has been to conceptualize and, when permitted, carry through on a cycle of evaluation, development, and evaluation. We expect, per Rossi's law, that many federal education programs are likely to be found to be ineffective in rigorous trials. We are designing our initial evaluations of those programs so that the results, if negative, will support hypotheses about how those programs might be strengthened. The second round of evaluation, in this model, involves funding development of new approaches or potential improvements to the program being delivered, and evaluating those supposed enhancements. The third round of evaluation involves disseminating those improved programs at scale and evaluating their effects. Thus the cycle is to evaluate, improve, extend, evaluate.

For example, we are following this model in our evaluation of after-school programs. The department's initial evaluation of the 21<sup>st</sup>-Century Community Learning Center's program found few if any positive effects for children attending those programs compared with children, chosen by lottery, whose parents were left to their own devices in obtaining after-school care (Office of the Undersecretary, 2002). One problem identified in the initial evaluation was a lack of education materials for after-school use that were engaging and consistent with the education tasks encountered by children during the regular school day. Thus the second round of evaluation is funding the development of new after-school curriculum materials in reading and math. The effects of these enhanced materials will be evaluated in a randomized trial. Given positive impact, the next step will be to encourage grantees to adopt the improved materials, followed by another evaluation of effects as these programs are routinely implemented in the field.

*Long timelines.* Most problems in education are immediate. Whether it be reducing achievement gaps, recruiting more qualified teachers, or selecting a mathematics curriculum, policy and practice can't wait. In areas in which research is weak, educators and policy-makers who might be willing to ground decisions in evidence are frequently frustrated to learn that the available evidence is meager





and that it will take several years to produce something that could support strong policy. Our response is to point out that research in all fields, including medicine, has long timelines. We also strongly encourage policymakers to adopt an experimental attitude when the research base is equivocal. For example, rather than take a gamble on a particular math curriculum for a district when the evidence on its effectiveness is weak, why not introduce it in a few schools and assess its impact before extending it to the whole district?

*Ideologues.* Many of the most passionate, involved advocates of education reform on the right and the left of the political spectrum are committed to particular ideas. On the left, funding adequacy and a highly professionalized teaching force are popular ideas. On the right, the value of competition and choice is a dominant philosophy. In the middle, standards and assessment-based reform are embraced. These are merely examples. There are dozens—if not hundreds—of ideas in play, most held strongly, and many argued well. Too many advocates of those ideas are interested only in evidence that supports their position. Disconfirming evidence is ignored or discounted. This is not unique to education. Indeed, it is characteristic of human thought and decision making. But unlike many other fields, in education the rules of evidence are in dispute, the institutional structures to support evidence-based decision making are weak, and the cultural stakes are very high. As a result, ideologues often have the upper hand. Our response at IES is to assiduously avoid being, or being perceived as, just another entity with an idea or opinion. We try to hew closely to what is actually known through rigorous empirical methods. In this way we hope to avoid attacks by ideologues who view us as taking sides with their enemy; at a minimum, we aspire to be blameless in the face of such attacks. At the same time, the whole enterprise I've described in this talk is intended to grow to a point where it can be a check on the fad and fancy that derives from ideologically driven policy. We value ideas about education and education reform, but in the context of the use of rigorous evidence to test those ideas. From this perspective, a good argument doesn't trump good evidence, and is no substitute for it.

## Summing Up

I've described some of the principles that guide our effort: that methods and relevance matter, and that we can learn from the experience of other fields that have already passed the tipping point in the swing from casual observation, intuition, and eminence to systematic empiricism. I've shared with you some of the practical advances we've made in making IES an organization that can produce, disseminate, and encourage the use of evidence in education. And I've characterized some of the formal and everyday political issues that play such an important role in the functioning of IES and advancing the mission of evidence-based education.



*C. Bradley Moore and Fay Lomax Cook thank Whitehurst (center) for his talk on evidence-based education.*

We need policymakers, educators, and concerned citizens to see the value of rigorous evidence, to turn to it when difficult decisions arise, and to insist that new policies that can't wait for evidence be tested as they are implemented. That requires a transformation in the way education is conducted. IES is at the center of that transformation because that is our statutory mission, we have more resources than anyone else to invest, and there is substantial, bipartisan political support for evidence-based education policy. This is a window of opportunity, not a guarantee of success. No matter how well IES does its job, the goal of evidence-based practice and policy in education will not be achieved unless there is an expanding coalition of researchers, practitioners, and policymakers who are willing to contribute to it with their time and talents. That coalition exists and is growing. The Institute for Policy Research at Northwestern University is very much a part of it. Thank you for doing what you do, and for inviting me to address you.



## References

- Cook, T. D., and D. T. Campbell. 1979. *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Co.
- Gardner, H. 2002. The quality and qualities of educational research. *Education Week*, 22, September 9: 72.
- Glazerman, S., D. M. Levy, and D. Myers, 2003. Nonexperimental versus experimental estimates of earnings impacts. *Annals of the American Academy of Political and Social Science*, 589:63-93.
- Institute of Education Sciences. 2003. *Institute of Education Sciences: Findings from interviews with education policymakers*. [www.ed.gov/rschstat/research/pubs/findingsreport.pdf](http://www.ed.gov/rschstat/research/pubs/findingsreport.pdf).
- National Research Council. 2004. *On evaluating curricular effectiveness: Judging the quality of K-12 mathematics evaluations*, by Committee for a Review of the Evaluation Data on the Effectiveness of NSF-Supported and Commercially Generated Mathematics Curriculum Materials. Washington, D.C.: National Academies Press.
- Office of the Undersecretary, U.S. Department of Education. 2002. *When schools stay open late: The national evaluation of the 21st-Century Community Learning Centers Program*. [www.ed.gov/pubs/21cent/firstyear/summ.html](http://www.ed.gov/pubs/21cent/firstyear/summ.html).
- Viadero, D. 2003. Study finds higher gains in states with high-stakes tests. *Education Week*, 22, April 16: 10.
- Viadero, D. 2002. Bill would remake OERI into 'Education Sciences' academy. *Education Week*, 21, March 6: 31.
- Winter, G. 2003. New ammunition for backers of do-or-die exams. *New York Times*, April 23: 9.
- Winter, G. 2002. Make-or-break exams grow, but big study doubts value. *New York Times*, December 28: 1.