At the dawn of the 21st century, educational research is finally entering the 20th century. The use of randomized experiments that transformed medicine, agriculture, and technology in the 20th century is now beginning to affect educational policy. This article discusses the promise and pitfalls of randomized and rigorously matched experiments as a basis for policy and practice in education. It concludes that a focus on rigorous experiments evaluating replicable programs and practices is essential to build confidence in educational research among policymakers and educators. However, new funding is needed for such experiments and there is still a need for correlational, descriptive, and other disciplined inquiry in education. Our children deserve the best educational programs, based on the most rigorous evidence we can provide.

Evidence-Based Education Policies:
Transforming Educational Practice and Research
by Robert E. Slavin

Education is on the brink of a scientific revolution that has the potential to profoundly transform policy, practice, and research. Consider the following:

- In 1998, Congress appropriated $150 million per year to provide schools funds to adopt "proven, comprehensive reform models" (U.S. Department of Education, 1999). This unprecedented legislation, introduced by Congressmen David Obey and John Porter, defined proven in terms of experimental-control comparisons on standards-based measures. To my knowledge, this was the first time in history that federal education funding has been linked directly to evidence of effectiveness (see Slavin, 1997). In 2001, Comprehensive School Reform (CSR) funding had been progressively increased to $310 million annually, and has provided funding to more than 2,600 mostly high-poverty schools (Southwest Educational Research Laboratory, 2002).

- Kent McGuire, Director of the Office of Educational Research and Improvement (OERI) in the Clinton administration, convinced Congress that the CSR funding program warranted a substantial increase in education research and development. Under his leadership, an array of capacity building, program development, and evaluation efforts were launched. All of these were intended to put programs with rigorous evidence of effectiveness into thousands more schools, particularly those serving many at-risk children.

- The Bush administration’s first domestic initiative, the reauthorization of the Elementary and Secondary Education Act (ESEA) called No Child Left Behind, took the idea of scientifically based practice to an even higher level. No Child Left Behind (U.S. Congress, 2001) mentions “scientifically based research” 110 times. It defines “scientifically based research” as “rigorous, systematic and objective procedures to obtain valid knowledge,” which includes research that “is evaluated using experimental or quasi-experimental designs,” preferably with random assignment. Scientifically based research is intended to serve as the basis for Title I programs, Reading First programs for reading in grades K–3, Early Reading First programs for pre-kindergarten, CSR, and many other components. Funding for ESEA overall was increased by 18%, the largest increase ever.

- Grover Whitehurst, the Bush Administration's OERI director, has taken a strong line in support of randomized experiments (Whitehurst, 2002). In a request for proposals that is a revolutionary document in its own right, OERI invited early childhood programs to subject themselves to randomized evaluations in which data would be collected by third-party evaluators (U.S. Department of Education, 2002a). Requests for proposals of this kind in other areas are in the pipeline. Not since the Follow Through Planned Variation studies of the 1970s (Rhine, 1981) have rigorous, experimental designs with common measures been applied to programs capable of broad scale replication.

- OERI, due to be reauthorized in 2002, will without any doubt be reorganized to focus resources on randomized and rigorous matched experimental research on programs and policies that are central to the education of large numbers of children. The U.S. Department of Education (2002b) strategic plan for 2002–2007 anticipates having 75% of all OERI–funded research that addresses causal questions use random assignment designs by 2004 (currently, such research probably represents less than 5% of causal research funded by OERI). As a direct result, Congress is likely to substantially increase funding for educational research.

It is important to note that none of these policy developments have yet produced the revolution I am anticipating. The CSR funding program, despite a clear focus on proven programs, has so far provided most of its funds to programs with little or no rigorous evidence of effectiveness, including many “programs” slapped together for the purpose of obtaining funding. A 1999 review of the research on 24 comprehensive reform models by the American Institutes of Research (AIR) (Herman, 1999) categorized them as having strong evidence of effectiveness; promising, marginal, mixed, weak, or no effects; or no research. Among
the definition of considered by AIR. The recent ESEA reauthorization tightens up unresearched, or to national or homegrown models not even grants (63.2%) have gone to programs either rated as mixed or strong evidence and 16.0% to programs rated as promising or marginal. Most of the 2,665 CSR grants made from 1998 to 2002 (Southwest Educational Research Laboratory, 2002), only 20.8% of grants have either rated as mixed or unresearched, or national or homegrown models not even considered by AIR. The recent ESEA reauthorization tightens up unproven and comprehensive, and places more emphasis on programs with scientifically based evidence of effectiveness (U.S. Department of Education, 2002c), but state officials who review CSR proposals still have broad discretion and could continue to minimize or ignore the research base behind the programs they fund. No Child Left Behind and other initiatives emphasizing rigorous research are too new to have had any impact on practice or funding. Yet these and other developments, if not yet proven, still create the potential for changes with far-reaching consequences. It is possible that these policy reforms could set in motion a process of research and development on programs and practices affecting children everywhere. This process could create the kind of progressive, systematic improvement over time that has characterized successful parts of our economy and society throughout the 20th century, in fields such as medicine, agriculture, transportation, and technology. In each of these fields, processes of development, rigorous evaluation, and dissemination have produced a pace of innovation and improvement that is unprecedented in history (see Shavelson & Towne, 2002). These innovations have transformed the world. Yet education has failed to embrace this dynamic, and as a result, education moves from fad to fad. Educational practice does change over time, but the change process more resembles the pendulum swings of taste characteristic of art or fashion (think hemlines) rather than the progressive improvements characteristic of science and technology (see Slavin, 1989).

Welcome to the 20th Century

At the dawn of the 21st century, education is finally being dragged, kicking and screaming, into the 20th century. The scientific revolution that utterly transformed medicine, agriculture, transportation, technology, and other fields early in the 20th century almost completely bypassed the field of education. If Rip Van Winkle had been a physician, a farmer, or an engineer, he would be unemployable if he awoke today. If he had been a good elementary school teacher in the 19th century, he would probably be a good elementary school teacher today. It is not that we have not learned anything since Rip Van Winkle’s time. It is that applications of the findings of educational research remain haphazard, and that evidence is respected only occasionally, and only if it happens to correspond to current educational or political fashions.

Early in the 20th century, the practice of medicine was at a similar point. For example, research had long since identified the importance of bacteria in disease, and by 1865 Joseph Lister had demonstrated the effectiveness of antiseptic procedures in surgery. In the 1890s, William Halsted at Johns Hopkins University introduced rubber gloves, gauze masks, and steam sterilization of surgical instruments and demonstrated the effectiveness of these procedures. Yet it took 30 years to convince tradition-bound physicians to use sterile procedures. If he dropped his scalpel, a physician in 1910 was as likely as not to give it a quick wipe and carry on.

Today, of course, the linkage between research and practice in medicine is so tight that no physician would dream of ignoring the findings of rigorous research. Because medical practice is so closely based on medical research, funding for medical research is vast, and advances in medicine take place at breathtaking speed. My father’s cardiologist recommended that he wait a few years to have a necessary heart valve operation because he was sure that within that short span of time, research would advance far enough to make the wait worthwhile. As it turned out, he was right.

The most important reason for the extraordinary advances in medicine, agriculture, and other fields is the acceptance by practitioners of evidence as the basis for practice. In particular, it is the randomized clinical trial—more than any single medical breakthrough—that has transformed medicine (Doll, 1998). In a randomized clinical trial, patients are assigned at random to receive one treatment or another, such as a drug or a placebo. Because of random assignment, it can be assumed with an adequate number of subjects that any differences seen in outcomes are due to the treatment, not to any extraneous factors. Replicated experiments of this kind can establish beyond any reasonable doubt the effectiveness (or lack thereof) of treatments intended for applied use (see Boruch, 1997).

Experiments in Education

In education, experiments are not uncommon, but they are usually brief, artificial experiments on topics of theoretical more than practical interest, often involving hapless college sophomores. Far more rare are experiments evaluating treatments of practical interest studied over a full school year or more. For example, there are many outstanding brief experiments published each year on the effects of various mnemonic teaching strategies. These have built a strong case for the effectiveness of mnemonic methods and detailed understanding of the conditions under which they work best (see Levin & Levin, 1990). However, I am unaware of any experiment that has evaluated, for example, a year-long course making extensive use of mnemonic devices. The research on mnemonic strategies is directly useful to teachers, who can be encouraged to teach, “When two vowels go walking, the first one does the talking,” or occasional mnemonics for remembering the order of planets out from the sun or trigonometric functions. Yet it is difficult to imagine that teaching and learning will make broad advances because teachers make occasional use of one or another mnemonic device. I write an educational psychology textbook (Slavin, 2003) that is full of research findings of this type, findings that are valuable in advancing theory and potentially valuable to teachers in understanding their craft. Yet the brief experiments, correlational studies, and descriptive studies that yield most of the information presented in my text or any other educational psychology text do not collectively add up to school reform. They are suggestions about how to think about daily teaching problems, not guides to the larger questions educators and policymakers must answer. Imagine that research in cardiology described heart function and carried out small-scale laboratory studies but never developed and tested an artificial heart valve. If this were the case, I would be an orphan.
Imagine that agricultural research studied plant growth and diseases but never developed and tested new disease-resistant crops. Educational research has produced many rigorous and meaningful studies of basic principles of practice but very few rigorous studies of programs and practices that could serve as a solid base for policy and practice and has had little respect for the studies of this kind that do exist. Because of this, policymakers have rarely seen the relevance of research to the decisions they have to make and therefore have provided minimal funding for research. This has led to a declining spiral, as inadequate investments in research lead to a dearth of the kind of large-scale, definitive research that policymakers would feel to be valuable, making these policymakers unwilling to invest in large-scale, definitive research.

**Shifting Policy Perspectives**

The dramatic changes in federal education policies referred to earlier could potentially reverse this declining spiral. If the new funding flowing into research can produce some notable successes, we could have an ascending spiral: rigorous research demonstrating positive effects of replicable programs on important student outcomes would lead to increased funding for such research, which would lead to more and better research and therefore more funding. More important, millions of children would benefit in the fairly near term. Once we establish replicable paradigms for development, rigorous evaluation, replication, and dissemination, these mechanisms could be applied to any educational intervention or policy. Imagine that there were programs under way all the time to develop, evaluate, and disseminate new programs in every subject and every grade level, as well as programs on school-to-work transitions, special education, gifted children, dropout prevention, English language learners, race relations, drug abuse prevention, violence prevention, and so on. Every one of these areas lends itself to a development—evaluation—dissemination paradigm, as would many more. Over time, each area would experience the step-by-step, irreversible progress characteristic of medicine and agriculture because innovations would be held to strict standards of evaluation before being recommended for wide scale use.

**Research Designs**

The scientific revolution in education will only take hold and produce its desired impacts if research in fact begins to focus on replicable programs and practices central to education policy and teaching, and if it in fact employs research methods that meet the highest standards of rigor. This begs an important question: What kinds of research are necessary to produce findings of sufficient rigor to justify faith in the meaning of their outcomes?

Recently, OERI’s director, Grover Whitehurst (2002) and other educational researchers (see, for example, Mosteller & Boruch, 2002) have been arguing that nothing less than randomized experiments will do for evaluations of educational interventions and policies. The strong emphasis on randomized experiments is welcome, but ironic. After many years of relative policy indifference to experiments of any kind, OERI is leaping over the rigorously matched experiment to demand randomized experiments.

The difference in the value of randomized and well-matched experiments relates primarily to the problem of selection bias. In a matched experiment, it is always possible that observed differences are due not to treatments but to the fact that one set of schools or teachers was willing to implement a given treatment while another was not, or that a given set of students selected themselves or were selected into a given treatment while others were not.

When selection bias is a possibility at the student level, there are few if any alternatives to random assignment, because unmeasured (often unmeasurable) pre-existing differences are highly likely to be alternative explanations for study findings. For example, consider studies of after-school or summer school programs. If a researcher simply compared students attending such programs to those not attending who were similar in pretest scores or demographic factors, it is very likely that unmeasured factors such as student motivation, parents’ support for education, or other consequential factors could explain any gains observed, because the more motivated children are more likely to show up. Similarly, studies comparing children assigned to gifted or special education programs to students with similar pretest scores are likely to miss key selection factors that were known to whoever assigned the students but not measured. If one child with an IQ of 130 is assigned to a gifted program and another with the same IQ is not, it is likely that the children differ in motivation, conscientiousness, or other factors. In these kinds of situations, use of random assignment from within a selected pool is essential.

In contrast, there are situations in which it is teachers or schools that elect to implement a given treatment, but there is no selection bias that relates to the children. For example, a researcher might want to compare the achievement gains of children in classes using cooperative learning, or schools using comprehensive reform models, to the gains made by demographically similar control groups starting at the same pretest levels. In such cases, random assignment of willing teachers or schools is still far preferable to matching, as matching leaves open the possibility that volunteer teachers or staffs are better than non-volunteers. However, the likely bias is much less than in the case of student self-selection. Aggregate pretest scores in an entire school, for example, should indicate how effective the current staff has been up to the present, so controlling for pretests in matched studies of existing schools or classes would control out much of the potential impact of having more willing teachers. For external validity, it is crucial to note that the findings of a well-matched experiment comparing volunteers to non-volunteers apply only to schools or teachers who volunteer, but the potential for bias is moderate (after controlling for pretests and demographic factors).

The importance of this discussion lies in the fact that randomized experiments of interventions applying to entire classrooms can be extremely difficult and expensive to do and are sometimes impossible. My colleagues and I at Johns Hopkins University are working with a third-party evaluator, the University of Chicago’s National Opinion Research Center, to do a randomized evaluation of Success for All, a comprehensive reform model. Recruiting schools for this study has been extremely difficult, even though we are offering substantial financial incentives to schools willing to be assigned at random to experimental or control groups. We initially offered schools $30,000 to par-
participate, a total of $1.8 million just for incentives. Yet this was not sufficient; we ultimately had to offer the program at no cost to schools (but at a cost of about $70,000 per school to the study). For the cost of doing this randomized study, we (and others) could have done two or three equally large-scale matched studies. It is at least arguable that replicated matched studies, done by different investigators in different places, might produce more valid and meaningful results than one definitive, once in a lifetime randomized study.

Still, fully recognizing the difficulties of randomized experiments, I think they are nevertheless possible in most areas of policy-relevant program evaluation, and whenever they are possible, they should be used. Reviews of research in other fields have found that matched studies generally find stronger outcomes than randomized studies, although usually in the same direction (e.g., Friedlander & Robins, 1995; Fraker & Maynard, 1987; Ioannidis et al., 2001). Four randomized experiments we have planned or executed at Johns Hopkins University and the Success for All Foundation illustrate the potential and the pitfalls. One of these, mentioned earlier, involves randomly assigning 60 schools to Success for All or control conditions for a 3-year experiment. Initially, we offered $30,000 to each school, but we got hardly any takers. Schools were either unwilling to take a chance on being assigned to the control for 3 years, or they could not afford the program costs beyond the $30,000 incentive. In spring 2002, we changed our offer. Schools willing to participate were randomly assigned to use Success for All either in Grades K–2 or in 3–5, at no cost. Recruitment was still difficult, but under this arrangement, we signed up adequate numbers of schools.

For another study proposed by my colleague Bette Chambers (but not funded), we recruited schools for a study of the Curiosity Corner preschool model. We offered schools the program for free, either in 2002—2003 or 2003—2004 (with random assignment to the two start dates). The 2003–2004 group would have served as the control group in 2002—2003. This delayed treatment control group design involving no cost to schools was easy for schools to accept, and we did not have serious recruiting problems.

In a study of an after-school tutoring program led by my colleague Toks Fashola, individual first graders whose parents agreed to have them participate were assigned at random to be tutored in spring 2002 or fall 2002. Again, the fall group served as a control group during the spring. Finally, my colleague Geoff Borman is doing randomized evaluations of summer school programs in which individual children are randomly assigned to participate now or later (Borman, Boulay, Kaplan, Rachuba, & Hewes, 2001). In these cases, obtaining sufficient volunteers was not difficult.

These examples of a diverse set of research problems illustrate that one way or another, it is usually possible to use random assignment to evaluate educational programs. There is no one formula for randomization, but with enough resources and cooperation from policymakers, random assignment is possible.

Beyond the benefits for reducing selection bias, there is an important political reason to prefer randomized over matched studies at this point in history. Because of political developments in Washington, we have a once-in-a-lifetime opportunity to reverse the “awful reputation” that educational research has among policymakers (Kaestle, 1993; Lagemann, 2002). This is a time when it makes sense to concentrate resources and energies on a set of randomized experiments of impeccable quality and clear policy importance to demonstrate that such studies can be done. Over the longer run, I believe that a mix of randomized and rigorous matched experiments evaluating educational interventions would be healthier than a steady diet of randomized experiments, but right now we need to establish the highest possible standard of evidence, on a par with standards in other fields, to demonstrate what educational research can accomplish.

On the topic of matched experiments, these, too, must be rigorous, planned in advance, and carefully designed to minimize selection bias. The hallmark of science is organized, disciplined inquiry that gives the null hypothesis every consideration. After the fact and pre-post experiments, for example, rarely meet this standard, but are all too common in program evaluations.

**Non-Experimental Research**

I should hasten to say that forms of research other than experiments, whether randomized or matched, can also be of great value. Correlational and descriptive research is essential in theory building and in suggesting variables worthy of inclusion in experiments. Our Success for All program, for example, owes a great deal to correlational and descriptive process-product studies of the 1970s and 1980s (see Slavin & Madden, 2001). As components of experiments, correlational and descriptive studies can also be essential in exploring variables that go beyond overall program impacts. In some policy contexts, experiments are impossible, and well-designed correlational or descriptive studies may be sufficient.

However, the experiment is the design of choice for studies that seek to make causal conclusions, and particularly for evaluations of educational innovations. Educators and policymakers legitimately ask, “If we implement Program X instead of Program Y, or instead of our current program, what will be the likely outcomes for children?” For questions posed in this way, there are few alternatives to well-designed experiments.

**Rigorously Evaluated Versus Programs and Practices Based on Scientific Research**

If the scientific revolution in education is to achieve its potential to transform education policy and practice, it must focus on research that is genuinely of high quality and is appropriate to inform the decisions that educators and policymakers face.

There is a key distinction that is easily lost in the current enthusiasm for science as a basis for practice. This is the distinction between programs and practices that are based on scientifically based research and those that have themselves been rigorously evaluated. In the No Child Left Behind legislation (U.S. Congress, 2001), the formulation, used 110 times, is “based on scientifically based research.” The difficulty here is twofold. First, any program can find some research that supports the principles it incorporates. The legislation and guidance for No Child Left Behind and Reading First specify in some detail the specific reading research findings expected to be implemented under federal funding, but it is still possible for a broad range of programs to claim to be based on scientifically based reading research.
More important, the fact that a program is based on scientific research does not mean that it is in fact effective. For example, imagine an instructional program whose materials are thoroughly based on scientific research, but which is so difficult to implement that in practice teachers do a poor job of it, or which is so boring that students do not pay attention, or which provides so little or such poor professional development that teachers do not change their instructional practices. Before the Wright brothers, many inventors launched airplanes that were based on exactly the same scientifically based aviation research as the Wright brothers used at Kitty Hawk, but the other airplanes never got off the ground.

Given the current state of research on replicable programs in education, it would be difficult to require that federal funds be limited to programs that have been rigorously evaluated, because there are so few such programs. However, programs that do have strong, rigorous evidence of effectiveness should be emphasized over those that are only based on valid principles, and there needs to be a strong effort to invest in development and evaluation of replicable programs in every area, so that eventually legislation can focus not on programs based on scientifically based research but on programs that have actually been successfully evaluated in rigorous experiments.

Research Syntheses
The evidence-based policy movement is by no means certain to succeed. Education has a long tradition of ignoring or even attacking rigorous research. Researchers themselves, even those who fundamentally agree on methodologies and basic principles, may disagree publicly about the findings of research. Individuals who oppose the entire concept of evidence-based reform will seize upon these disagreements, which are a healthy and necessary part of the scientific process, as indications that even the experts disagree.

For these and many other reasons, it is essential that independent review commissions representing diverse viewpoints be frequently constituted to review the research and produce consensus on what works, in language that all educators can access. In the area of reading, it is impossible to overstake the policy impact of the National Reading Council (Snow, Burns, & Griffin, 1998) and National Reading Panel (1999) reports, which produced remarkable consensus on the state of the evidence. Consensus panels of this kind, with deep and talented staff support, should be in continual operation on a broad range of policy-relevant questions so that practitioners and policymakers can cut through all the competing claims and isolated research findings to get to the big picture findings that methodologically sophisticated researchers agree represent the evidence fairly and completely.

Evidence in the Age of Accountability
Evidence-based policies for education would be important at any time, but they are especially important today, given the rise of accountability. State and national governments are asserting stronger control over local education, primarily by establishing consequences for schools based on gains or losses on state assessments. The accountability movement is hardly new; it has been the dominant education policy focus since the early 1980s. Accountability systems are becoming more sophisticated and are reaching further into educational practice, but accountability is still a necessary but insufficient strategy for school reform. Most obviously, teachers and administrators need professional development, effective materials, and other supports closely aligned with state standards to help them move from where they are to where they need to be.

However, rewards and sanctions based on test score gains can be very inexact in fostering good practice. Year-to-year changes in an individual school are unreliable indicators of a school’s quality (see Linn & Haug, 2002). Scores can fluctuate on a year-to-year basis for a hundred reasons that have nothing to do with program effectiveness. These include population changes, student mobility, changes in special education or bilingual policies, test preparation, and changes in promotion policies, as well as random factors. Also, it takes so long to report test scores that schools may be into the following school year still doing the wrong thing before they find out.

For these and other reasons, it is essential that schools focus both on the evidence base for their programs and on the outcomes in their particular school. Hospitals might be held accountable for their success rates with various conditions, but they would never dream of implementing procedures discordant with rigorous, widely accepted research. Similarly, schools should be expected to use methods known to be effective in general and then to make certain that in their particular school implementation of those methods is of sufficient quality to ensure progress on the state assessments.

Will Educational Research Produce Breakthroughs?
In a recent op-ed in Education Week, James Gallagher (2002) argued that evidence-based policies are setting up false expectations among policymakers. He maintained that unlike medical research, educational research is unlikely to produce breakthroughs. He is perhaps right about this; education reform is not apt to invent exotic treatments that have the immediate impact that the Salk vaccine had on polio. However, the value of evidence-based policies does not depend on breakthroughs. Fields that invest in research and development often produce progressive, step-by-step improvements. Modern automobiles use the internal combustion engine, just like the Model T, but modern automobiles are far more efficient and effective. Physicians could remove ruptured appendices at the turn of the century, but these operations are far less risky now. In these and hundreds of other cases, it is the accumulation of small advances rather than breakthroughs that led to substantially improved practice. This is how evidence-based policies will probably improve education. Once we have dozens or hundreds of randomized or carefully matched experiments going on each year on all aspects of educational practice, we will begin to make steady, irreversible progress. Until then, we are merely riding the pendulum of educational fashion.

Potential Impact of Evidence-Based Policies on Educational Research
Up to now, I have written primarily about the potential impact of evidence-based policies on education policies and practice. I
would now like to consider the potential impact on educational research. I believe that if evidence-based policies take hold, this will be enormously beneficial for all of educational research, not just research involving randomized or matched experiments. First, I am confident that when policymakers perceive that educational research and development is actually producing programs that are shown in rigorous experiments to improve student outcomes, they will fund research at far higher levels. This should not be a zero-sum game in which new funds for experiments will be taken from the very limited funds now available for educational research (see Shavelson & Towne, 2002). Rather, I believe that making research relevant and important to policymakers will make them more, not less, willing to invest in all forms of disciplined inquiry in education, be it correlational, descriptive, ethnographic, or otherwise. The popularity of medical research depends totally on its ability to cure or prevent diseases, but because randomized experiments routinely identify effective treatments (and protect us from ineffective treatments), there is vast funding for basic research in medicine, including epidemiological, correlational, and descriptive studies. Researchers and developers will be able to argue convincingly that basic research is essential to tell us what kinds of educational programs are worth evaluating.

A climate favorable to evidence-based reform will be one in which individual researchers working on basic problems of teaching and learning will be encouraged and funded to take their findings from the laboratory or the small-scale experiment—or from the observation or interview protocol—to develop and then rigorously evaluate educational treatments. Education is an applied field. Research in education should ultimately have something to do with improving outcomes for children.

Conclusion

Evidence-based policies have great potential to transform the practice of education, as well as research in education. Evidence-based policies could finally set education on the path toward the kind of progressive improvement that most successful parts of our economy and society embarked upon a century ago. With a robust research and development enterprise and government policies demanding solid evidence of effectiveness behind programs and practices in our schools, we could see genuine, generational progress instead of the usual pendulum swings of opinion and fashion.

This is an exciting time for educational research and reform. We have an unprecedented opportunity to make research matter and to then establish once and for all the importance of consistent and liberal support for high-quality research. Whatever their methodological or political orientations, educational researchers should support the movement toward evidence-based policies and then set to work generating the evidence that will be needed to create the schools our children deserve.

NOTE

Dewitt Wallace-Reader’s Digest Distinguished Lecture, presented at the annual meeting of the American Educational Research Association, New Orleans, April 2, 2002. This article was written under funding from the Office of Educational Research and Improvement, U.S. Department of Education (Grant No. OERI-R-117-D40005). However, any opinions expressed are those of the author and do not necessarily represent OERI positions or policies.

REFERENCES


AUTHOR

ROBERT E. SLAVIN is Co-Director of the Center for Research on the Education of Students Placed at Risk at Johns Hopkins University, 200 W. Towson Boulevard, Baltimore, MD 21204-5200; rslavin@SuccessForAll.net. His research interests include comprehensive school reform, cooperative learning, evidence-based policy, and school and classroom organization.

Manuscript received May 7, 2002
Revisions received June 24, 2002
Accepted July 17, 2002