



# Education Research Can and Must Address “What Works” Questions

by Robert E. Slavin

Every year, teachers, principals, superintendents, and other educators have to make hundreds of decisions of potentially great importance to students. What reading program is most likely to ensure high reading performance? Should “social promotion” be used or should retentions be increased? Should summer school programs be provided to struggling students? Should kindergarten be half day or full day? What are the most effective means of providing remediation to children who are falling behind? In a word: What works?

In his response to my article on evidence-based education policies (Slavin, 2002), David Olson (2004) feels that education research has little to offer educators who face such decisions, and fundamentally cannot offer very much. Each of these “treatments,” he believes, is too diverse, too context bound, to permit useful generalizations. It is surely true that there are no “pure” treatments in education and that effects vary according to contexts and vary from child to child and school to school. Yet recognizing this variation, is it impossible to tell a teacher, principal, or superintendent anything at all about the likely average effects of one or another program or practice?

If one believes, as I do, that research in education has an obligation to answer the “what works” questions that educators, parents, and policymakers ask, then our job is to produce answers that are well justified. Olson’s simplistic characterization of experimental research pretends that such research can only produce *x* versus *y* comparisons, yet he surely knows that well-designed studies also characterize the conditions under which *x* works better or worse than *y*, identify the students for whom *x* works better or worse than *y*, and often produce rich qualitative information to supplement the quantitative comparisons. Such experiments (whether randomized or well matched) are admittedly rare, but that’s exactly the point of the recent policy shifts—to make them less rare.

Olson rejects the analogy to medical research made by many advocates of randomized experiments, because educational treatments are not a “pill” that is the same each time it is administered. Again, this is a simplistic characterization of medical research. Very often, medical treatments are quite complex and their impacts can vary depending on many factors. For example, a heart bypass operation involves many medical practitioners who may do their jobs well or poorly. Bypass operations may be differentially appropriate for patients of different diagnoses, ages, genders, or conditions. Double blind studies are impossible with

operations. Yet randomized evaluations of complex medical treatments are routinely done, and they establish with a high degree of confidence the effectiveness of given treatments under given circumstances for given patients. There is no fundamental reason that research in education cannot do the same.

One key assumption in Olson’s response is that there have been thousands of experiments evaluating educational practices and that they haven’t amounted to much. There may be thousands of brief laboratory studies, but experimental studies of replicable treatments that take place over a semester or more are extremely rare. I examined the *American Educational Research Journal (AERJ)* over the period 2000–2003 (Volumes 37–39 and Volume 40, No. 1). Out of 112 articles, six (5.4%) reported experimental-control comparisons. Of these, two were brief (4 weeks and 2 hours, respectively), and one used a matched design, leaving three (2.7%) randomized studies of extended treatments. A tenfold increase in randomized experiments of significant duration, to 27% of articles in *AERJ*, would still leave 73% of studies in *AERJ* using non-experimental methods (or brief experiments). In a recent review of research on bilingual vs. immersion elementary reading programs for English-language learners, my colleague and I located only three randomized experiments and 14 matched experiments that met minimal criteria for methodological adequacy (Slavin & Cheung, 2003). This is a slim basis for decisions on such an important question. Similarly small numbers of experiments have been done on most educational policies and practices of great importance. We need many more such studies, so that we can give practitioners and policymakers reliable, well-founded information on the decisions they must make.

The ultimate beneficiaries of education research must be children, not the researchers themselves. Enlightened educators look to education research for well-founded evidence to help them do a better job with the children they serve. For many questions, nonexperimental methods are perfectly appropriate. However, for questions that compare the outcomes of alternative programs or policies, there is no substitute for a well-designed experiment. Educators as well as researchers should celebrate the administration’s policy promoting such experiments and using their results to help improve outcomes for children.

## NOTE

This article was written under funding from the Institute of Education Sciences (IES), U.S. Department of Education (Grant No. R-117-D40005). However, any opinions expressed are those of the author and do not necessarily represent IES positions or policies.

The author would like to thank Sam Stringfield, Nancy Madden, Bette Chambers, and Cheryl Sattler for their comments on an earlier draft of this article.

## REFERENCES

- Olson, D. (2004). The triumph of hope over experience in the search for "what works": A response to Slavin. *Educational Researcher*, 33(1), 24–26.
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15–21.
- Slavin, R. E., & Cheung, A. (2003). *Effective programs for English language learners: A best-evidence synthesis*. Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

## AUTHOR

ROBERT E. SLAVIN, Co-Director of the Center for Research on the Education of Students Placed at Risk at Johns Hopkins University and Chairman of the Success for All Foundation, 200 W. Towsontown Blvd., Baltimore, MD 21204; rslavin@successforall.net. His research interests include comprehensive school reform, cooperative learning, school organization, at-risk students, and research review.

Manuscript received August 26, 2003

Final revision received August 26, 2003

Accepted September 29, 2003

### CALL FOR NOMINATIONS

*The American Educational Research Association*

#### WILLYSTINE GOODSSELL AWARD

To recognize an Educator who has served AERA on Behalf of Women, Girls, and Education through:

- Scholarship
- Activism
- Community building

This award will be presented by the AERA SIG Research on Women and Education in cooperation with SAGE and Women Educators at AERA's Annual Meeting in April 2004.

Send letters of nomination (electronically or via regular mail), including résumés of nominees, by March 1, 2004, to:

**Dr. Beverly J. Irby**  
**Sam Houston State University**  
**PO Box 2119**  
**Huntsville, TX 77341**  
**E-mail: edu\_bid@shsu.edu**  
**Phone: 936-294-1134**

Among the list of past **WILLYSTINE GOODSSELL AWARDees** are Elizabeth Fennema, Sari Knopp Biklen, Susan Bailey, Selma Greenberg, Maxine Greene, Susan Klein, Carol Shakeshaft, Diane Pollard, Nel Noddings, Jane Roland Martin, Patricia Campbell, Marcia Linn, Sandra Hollingsworth, Gwen Baker, Pat Schmuck, and Catherine Marshall.