

Evidence-Based Practice

**Evidence-Based Practice:
Building Capacity for Informed
Professional Judgment**
Warren Simmons

**Bringing Measurement to
District-Based Accountability:
The Challenge for State
Education Departments**
David V. Abbott

**When Districts Use Evidence
to Improve Instruction:
What Do We Know and
Where Do We Go from Here?**
Meredith I. Honig and Cynthia E. Coburn

**Data, Observations, and
Tough Questions: A Journalist's
Evidence of School Quality**
Dale Mezzacappa

Voices in Urban Education
Number 6, Winter 2005

Publisher
Annenberg Institute for School Reform
at Brown University
Warren Simmons, Executive Director

Editor
Robert Rothman

Managing Editor
Susan Fisher

Copy Editor
Margaret Balch-Gonzalez

Production & Distribution
Mary Arkins

Design
Gilbert Design Associates, Inc.

Illustration
Carolina Arentsen

Voices in Urban Education (ISSN 1553-54X) is published quarterly at Brown University, Providence, Rhode Island. Articles may be reproduced with appropriate credit to the Annenberg Institute. Single copies are \$10 each, including postage and handling. A discount is available on bulk orders. Call 401 863-2018 for further information.

The Annenberg Institute for School Reform was established in 1993 at Brown University. Its mission is to develop, share, and act on knowledge that improves the conditions and outcomes of schooling in America, especially in urban communities and in schools serving disadvantaged children. For program information, contact:

Annenberg Institute for School Reform
Brown University, Box 1985
Providence, Rhode Island 02912
Tel: 401 863-7990
Fax: 401 863-1290
Web: www.annenberginstitute.org

©2005 Annenberg Institute for School Reform

Evidence-Based Practice: Building Capacity for Informed Professional Judgment

Warren Simmons

The current emphasis in education reform is on scientific research that answers the question “What works?” However, this approach often fails to provide useful guidance for school practitioners. To improve student outcomes, researchers and practitioners must collaborate on research that explains why a particular practice meets students’ needs and what conditions are needed to enable the practice to work.

Education reforms over the past two decades have tended to emphasize *will* rather than *skill*. The assumptions behind standards-based reform seem to be that educators know what to do, but that they lack the incentives and the flexibility to take appropriate steps to improve performance. As Richard Elmore and Robert Rothman (1999) observed, the implied theory of action underlying standards-based reform held that if schools were given more resources and flexibility in exchange for being held accountable for getting all students to meet high standards, teachers and principals would have the motivation needed to foster continuous improvement in student achievement.

While the prevalence of school-funding litigation casts doubt on whether resources are adequate for improvement, serious questions have also been raised about whether or not practitioners – particularly, those in urban settings – have the knowledge needed to improve conditions of learning and thereby raise student outcomes. Most notably, highly qualified teachers are in short supply. Dysfunctional human resource systems and

chronic funding shortfalls leave many urban school districts poorly prepared to compete for skilled teachers. These same human resource systems also operate to assign novice or struggling teachers to urban schools serving students with the greatest academic and economic needs (DeStefano & Foley 2003).

Research as a Capacity-Building Tool

Standards-based reform, however, has not been entirely blind to the need for greater expertise to guide school and classroom improvement. America 2000, the education reform proposed by President George Bush in 1990, created the New American Schools (NAS) Development Corporation to sponsor a set of research-based whole-school reform designs that schools could adopt to help students reach high standards.

This approach assumed that schools lacking the internal expertise to promote student achievement could be improved significantly by acquiring a carefully articulated design that was research based. In 1998, Congress affirmed this reasoning by passing legislation authorizing \$150 million to pro-

Warren Simmons is executive director of the Annenberg Institute for School Reform.

Research-based designs or programs, by themselves, do not guarantee improvement if conditions supporting their effective use are not present.

vide schools with the resources needed to implement what were then called *research-based comprehensive school reform designs*. By this time, the eleven original models created by NAS had been joined by over forty designs possessing various forms of research that demonstrated their efficacy and/or origins in a particular line of research.

The term *research-based*, as applied to designs, has two separate meanings. One has to do with the degree to which a design and its various elements are rooted in research on learning, development, and other areas related to education. The second involves the degree to which a particular design has evaluation data demonstrating its effectiveness. The models listed as being research-based vary along these two dimensions of research considerably, yet they all carry the federal government's stamp of approval.

The weight placed on research-based school designs as a solution for failing schools rests on the quality of available research, on the one hand, and the capacity of schools and districts

to implement a design faithfully and effectively, on the other. In a report evaluating the progress of the eleven original NAS school designs, Thomas Glennan (1998) observed that the "success" of a design in a particular school was a function of several factors:

The performance of a school would clearly be a function of the design, the assistance received as it was implemented, the nature of the school itself, and a variety of qualities of a district's operating environment. Attributing a school outcome to the design is clearly inappropriate. (p. 79)

In short, research-based designs or programs, by themselves, do not guarantee improvement if conditions supporting their effective use are not present. Charles Payne and Mariame Kaba (2001) made similar observations about the difficulty low-capacity schools in Chicago encountered in making effective use of research-based school reform designs that required a baseline set of conditions (e.g., collaborative culture, shared vision) that were often weak or absent in chronically failing schools. Moreover, the conditions that predict improvement include a school's ability to match its needs with the elements of a design, the ability of an external provider to offer adequate professional development and support services, and the capacity of the district governing the school to align its supports and services with those provided by the design (Hassel 2002).

This cautionary tale about the transferability of research-based designs and programs has not dampened belief in the power of research to inform practice and improve schools. If anything, No Child Left Behind (NCLB), the 2001 Reauthorization of Title I of the Elementary and Secondary Education Act, has amplified the importance of research by making the receipt of

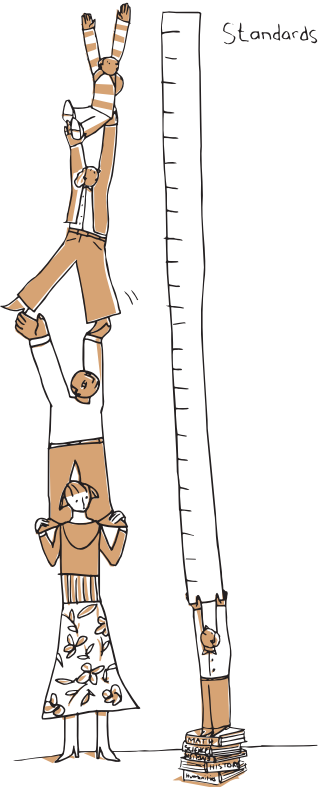
Title I dollars contingent on schools' and entire districts' adopting research-based curricular programs, along with offering incentives for the use of research-based school designs. Moreover, NCLB has expanded the importance of research, while narrowing the definition of acceptable research to studies using randomized assignment of subjects to experimental and control groups. This restriction is based on the largely unsubstantiated assumption that the weak connection between research on what we know and what we are able to achieve in practice has had more to do with the quality of existing research than with the conditions affecting its use in practice.

NCLB and the U.S. Department of Education's new Institute of Education Sciences attempt to address the "quality" problem by creating strict scientific standards for research supported by the department and by tying federal Title I dollars to the use of programs or designs documented by research that meets these standards.

Expanding Available Evidence

NCLB also bolsters research or evidence-based practice in another way. The law requires states to issue school and district reports that reveal whether schools and their major student subgroups – African Americans, Hispanics, etc. – are making adequate yearly progress toward proficiency in mathematics and reading.

This federal requirement, in effect, establishes a base of data that educators and the public can use to inform their judgments about the needs of particular schools and districts. Presumably, the widespread availability of data on the status and needs of students, schools, and districts, coupled with more robust research-based curricular programs and school designs, should



serve to accelerate the impact of research on practice and foster improvements in school and student performance.

Self-Imposed Limitations

NCLB should be applauded for stressing the importance of research and of making disaggregated data on student performance at the school and district levels widely available and publicly reported. But the potential impact of these measures is undermined by the narrow definition of research and limited scope of the data covered by the law.

NCLB's reliance on standardized tests as a primary source of data to inform public judgments about student and school performance has been criticized widely. These measures typically assess a narrow range of desired skills and knowledge and provide little information about the conditions of learning that produce performance disparities. They often cause schools to narrow the

curriculum in a desperate effort to raise test performance (Herman 2003). At the same time, the law's emphasis on randomized trials provides little incentive for states and districts to invest in other types of research that might be more valuable and could inform practice in more useful ways.

Monitoring Conditions of Instruction

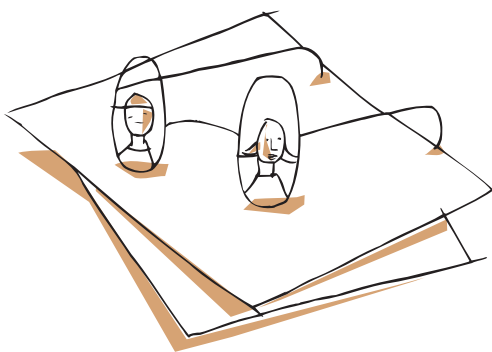
The shortcomings of standardized tests are well known, but alternatives that would provide broader and more useful information and lessen the possibility of perverse effects are less familiar.

The National Research Council's Title I Assessment Committee argued that schools and districts should be aided in making the links between assessment, learning, and practice by developing measures that monitor the conditions of instruction. These kinds of measures would assess the quality of the *enacted curriculum* – shorthand for what Richard Elmore (2002) describes as the intellectual and material resources that influence the interactions of teachers and students with content (e.g., subject-matter knowledge combined with students' prior knowledge).

To gauge the quality of the enacted curriculum, schools, school districts, and, to some extent, states would have to invest in the systematic and ongoing use of *formative assessments*. These assessments are needed to shed light on the alignment of state and district standards with the intellectual rigor and quality of

- instructional texts and materials;
- teacher lessons and assignments;
- classroom discourse between and among students and teachers; and
- student work.

Models of such formative assessments exist and include Lesson Study,



an approach modeled after teaching practices in Japan, and the Teaching and Learning Review, a method developed by the Annenberg Institute for School Reform to assess the match between the *intended curriculum* and the enacted curriculum. To date, though, these and other models for ongoing formative assessment remain confined to networks of teachers and schools. To become tools that bridge the gap between standards and achievement, they must be elevated to become part of a system of assessments used by schools and districts to improve instruction in the manner described by Elmore.

Aligning Clinical and Research Knowledge

The development and systematic use of formative assessments – focused on conditions of instruction in schools and the relation of these conditions to district policies and supports – would provide practitioners the data they need to guide their professional judgment. Formative assessments of classroom, school, and district practice would also help practitioners weigh the appropriateness and relevance of basic and evaluation research touting specific programs and practices.

The current emphasis on scientific research that demonstrates “what works” overlooks a second critical question for the practitioner: What works, given the needs and values of my students and community and the condition and capacity of my school and district? If this second question is ignored, schools can be led to choose research-based designs and programs that don’t address the needs of their learners and practitioners. For instance, many of the research-based whole-school reform designs lack supports for English-language learners and students with disabilities, yet they

The current emphasis on scientific research that demonstrates “what works” overlooks a second critical question for the practitioner: What works, given the needs and values of my students and community and the condition and capacity of my school and district?

carry the U.S. Department of Education’s seal of approval as being “proven.” In this case, it is left up to the “buyer” to decide if *proven* translates into *effective* for the constellation of teachers and students in a particular school or school district.

Unfortunately, the context-specific nature of effective practice is overshadowed by the implication of the “what works” incantation that highlights results but obscures the conditions needed to ensure effectiveness. To advance practice and improve student outcomes, researchers and practitioners must collaborate to conduct research that addresses five basic questions:

- What are the needs of the students being served and the practitioners charged with supporting their learning?



- What are the conditions of instruction within and across challenged schools or schools needing improvement?
- How do district policies and supports work to affect these conditions for better or worse?
- Given answers to the first three questions, which research-based programs or designs are best suited to addressing the needs and conditions present in my school/district?
- What are the short-term, intermediate, and long-term resources and strategies I must consider to ensure effective implementation of a particular program or design?

These questions are rarely answered by basic or evaluation research because the design of such studies pays more attention to whether an effect occurred or not, as opposed to how the effects hold up under a range of conditions. It reminds me of my youthful rush to buy a BMW, the “ultimate driving machine.” It might have been the ultimate driving machine in California, but in the hilly and icy driving conditions in Ithaca, New York, where I attended graduate school, it became a “near-death driving machine” that I left parked in the garage for much of the time.

Unfortunately, schools and districts often adopt effective programs and practices that wind up like my BMW, parked in the garage due to a poor match between the needs of the school or district and the components of a particular program or design (Corcoran, Fuhrman, and Belcher 2001).

Integrating Research and Practice

Scientific research alone will not build a stronger knowledge base about the relevance of research-based educational programs and designs to specific contexts found in urban schools. Research

addressing this issue must be sensitive to the challenges and conditions in urban schools and make this variability a focal point of study rather than a factor to be controlled. Denis Newman (2004) makes a similar point in a recent critique of the emphasis on scientific research in education:

An instructional program cannot be “proven effective” by a single program of research. Unless decision-makers can recognize their own situation in the site where the research was conducted, they have little reason to think that the same effect will occur for them. The best evidence will always be that collected locally. (p. 9)

Rather than dismissing the importance of scientific research, Newman and others argue for the need to augment randomized, controlled experiments with what Fritz Staub (2004) calls “design experiments.” According to Staub:

In design experiments, researchers engage in close collaboration with practitioners, and they are jointly accountable for the experiments carried out. Design-teams consisting of practitioners and researchers allow detailed local knowledge from the field of practice to be included in the design process, and the close collaboration is instrumental in ensuring that interventions are implemented as planned. Design experiments are to develop theories as well as new forms of practice through repeated cycles of designing, implementation, and analyses. (p. 44)

I would only depart from Staub’s position by giving priority to generating new forms of practice over developing new theories, given the moral imperative of school improvement. The kinds of research Staub describes are now being carried out by a new generation of locally focused, applied-research

organizations such as Research for Action in Philadelphia, Education Matters in Boston, and the Center for Research on the Context of Teaching at Stanford University. Typically, these organizations define problems and conduct research in collaboration with teams of researchers and practitioners.

This kind of close relationship helps bridge the gap between research and practice, as well as between research and clinical knowledge, and ensures that research is attuned to local conditions. These kinds of collaborations also ensure that practitioners’ concerns with *how* a program works are given equal weight with researchers’ emphasis on describing *what* works. Moreover, this

These kinds of collaborations also ensure that practitioners’ concerns with *how* a program works are given equal weight with researchers’ emphasis on describing *what* works.

approach circumvents the need to figure out how to translate research into practice by involving practitioners in the research itself. As a result, practitioners serve as researchers and translators simultaneously.

While this type of research has the potential to quicken the pace of learning, development, and improvement, the conditions and incentives needed to advance its development bear some examining. If school and central office practitioners are to be continually involved in these kinds of endeavors, district structures and policies must support collaborative inquiry and provide

the time and incentives needed to build partnerships with researchers.

Districts provide fertile ground for practitioner collaboration when they support the development of embedded professional learning communities (groups of practitioners working within and across schools, areas or regions of a district, and the district central office) with the features outlined by the Annenberg Institute (2004) and by Milbrey McLaughlin and Joan Talbert (2003). This fertile ground is further enriched by an emphasis on evidence-based practice, tools, and strategies that support the collection and sharing of data,

Researcher-practitioner collaboration also requires changes in the incentives and structure of higher education.

The current reward structure for research in higher education values contributions to the field over contributions to practice.

a shared vision of instruction, strong internal accountability, time for planning and reflection, distributed leadership, and continuous cycles of inquiry.

Researcher-practitioner collaboration also requires changes in the incentives and structure of higher education. The current reward structure for research in higher education values contributions to the field over contributions to practice. Junior faculty facing tenure decisions must publish in refereed journals where scholars determine the merits of the research based on the extent to which it contributes to discipline-based theory. Design experiments, in

contrast, are often interdisciplinary in nature and are oriented toward practice and product rather than driven by theory. Recent efforts by some universities to create clinical professorships reflect a nascent but significant recognition by a segment of higher education of the central importance of this kind of research to the mission of the university.

As “evidence-based practice” becomes a new watchword in education reform, its potential to improve conditions and outcomes in urban schools, particularly those serving disadvantaged children and youths, depends on how we understand and use it. We must be able to grasp the “deep structure” needed to attain meaningful evidence-based practice. We must also be able to advocate for the changes required in K–12 and higher education to make this work a valued and common expectation for researchers and practitioners. If we fall short in these tasks, evidence-based practice will find itself “garaged” with other promising reforms of the past.

References

- Annenberg Institute for School Reform. 2004. *Professional Learning Communities: Professional Development Strategies that Improve Instruction*. Providence, RI: Annenberg Institute for School Reform, Brown University. Available for download at <www.annenberginstitute.org/publications/plc.html>.
- Corcoran, T., S. H. Fuhrman, and C. L. Belcher. 2001. “The District Role in Instructional Improvement,” *Phi Delta Kappan* 83, no. 1:78–84.
- DeStefano, J., and E. Foley. 2003. “The Human-Resource Factor: Getting – and Keeping – Good Teachers in Urban Districts,” *Education Week* (April 16). Available online at <www.edweek.org/ew/articles/2003/04/16/31destefano.h22.html>.
- Elmore, R. F. 2002. *Bridging the Gap between Standards and Achievement: The Imperative for Professional Development in Education*. Washington, DC: Albert Shanker Institute.
- Elmore, R., and R. Rothman, eds. 1999. *Testing, Teaching, and Learning: A Guide for States and School Districts*. Washington, DC: National Academy Press.

Glennan, T. 1998. *New American Schools After Six Years*. Santa Monica, CA: RAND.

Hassel, B. 2002. *Making Good Choices: A Guide for Schools and Districts*. Revised edition, updated by Maria Ferguson. Naperville, IL: North Central Regional Educational Laboratory. Available online at <www.ncrel.org/csri/choices/makegood/title.htm>.

Herman, J. L. 2003. "State Test Lessons: The Effects on School Reform," *Voices in Urban Education* 1 (Spring):46–55. Providence, RI: Annenberg Institute for School Reform, Brown University.

McLaughlin, M., and J. Talbert. 2003. *Reforming Districts: How Districts Support School Reform*. Seattle: Center for the Study of Teaching and Policy, University of Washington.

Newman, D. 2004. *Generating the Scientific Evidence Schools Need so as to Leave No Child Behind*. Palo Alto, CA: Empirical Education Inc.

Payne, C., and M. Kaba. 2001. "So Much Reform, So Little Change: Building-Level Obstacles to Urban School Reform." Unpublished manuscript, Duke University.

Staub, F. C. 2004. "Transforming Educational Theory into Usable Knowledge: A Case of Co-constructing Tools for Lesson Design and Reflection." In *Quality in Practice-Oriented Research in Science Education*, edited by B. Ralle and I. Eilks, pp. 41–52. Aachen, Germany: Shaker.