Scientific Research and Evidence-Based Practice

Paul D. Hood

July 1, 2003

WestEd®

Improving education through research, development, and service

Table of Contents

Intent of this Paper

For the past decade education has been among the top agenda issues at national and state levels. Along with calls for higher education standards and accountability have emerged concerns that education research should play a more significant role in supporting education reform. The No Child Left Behind Act of 2001 is remarkable in its more than one hundred references to "scientifically-based research." More recently, the Institute of Education Sciences (IES) has promulgated the concept of "evidence-based education." In this paper we examine both of these concepts, place them in a larger context, and discuss their implications.

Because evidence-based education (EBE) is a newer and less familiar concept than scientific research and because EBE may have more immediate and pervasive implications, we examine it more closely. Over the past decade, evidence-based practice in the field of medicine has emerged as the model on which EBE seems to be based. Therefore, we examine this medical model in detail, compare it with EBE, and use it as a basis for projecting likely opportunities for development of EBE products and services. But first, to place this discussion in perspective, we briefly review concepts of research and development and examine various related definitions.

Conceptions of Research and Development

BASIC AND APPLIED RESEARCH

Vannevar Bush laid the basis for post-World War II American science policy. In "Science The Endless Frontier: A Report to the President on a Program for Postwar Scientific Research," Bush (1945) laid out an ambitious agenda. Central to Bush's view of science was a clear distinction between basic and applied research. According to Bush, "basic research is performed without thought of practical ends. It results in general knowledge and an understanding of nature and its laws." It creates the "scientific capital" and "the fund" for subsequent applied research. These distinctions between basic and applied research have continued in use for the past half-century.

As pointed out by Donald Stokes (1994, 1997) this sharp distinction between basic and applied research was made deliberately to provide the justification to ensure federal support for basic research following the conclusion of World War II. Bush's conception was deeply influential not only in the design of federal scientific agencies, most notably the National Science Foundation, but also in academic thinking.

Witness these recent National Science Foundation definitions

- Basic research is defined as research directed toward increases in knowledge or understanding of the fundamental aspects of phenomena and of observable facts without specific application toward processes or products in mind.
- Applied research is defined as research directed toward gaining knowledge or understanding necessary for determining the means by which a recognized and specific need may be met.
- Development is the systematic use of the knowledge or understanding gained from research directed toward the production of useful materials, devices, systems or methods, including design and development of prototypes and processes. (NSF, National Patterns of R&D Resources, 1998)

Bush's conception of science extended to virtually all disciplines and fields, including research in education where in reports on the state of education research dating from the 1950s, there were calls for strong basic research to be followed by applied research, development, and efforts at application. However, by the late 1960s, Cronbach and Suppes offered a somewhat different interpretation.

DISCIPLINED INQUIRY IN EDUCATION

When Cronbach and Suppes (1969) produced a report on the role of research in the improvement of education, they distinguished between two kinds of research: "conclusion-oriented research," which is concerned with testing hypotheses and developing theory, and "decision-oriented research," which deals with making education choices. They subsumed both kinds of research under the rubric "disciplined inquiry." Cronbach and Suppes argued that what set disciplined inquiry apart from other forms of inquiry was a number of characteristics, including:

1. Meaningful topics are addressed;
2. Systematic, clearly described procedures are employed and described so that readers can follow the logic of the study and assess the validity of the study's conclusions;
3. There is a sensitivity to the errors that are associated with the methods employed and efforts are made to control the errors or consider how they influence the results;
4. Empirical verification and sound logic are valued; and
5. Plausible alternative explanations are considered.

Thus disciplined inquiry, especially conclusion-oriented, is carried out in such a way that arguments can be examined in detail. The difference between the two approaches is mainly determined by the purposes of the research: to derive conclusions that would contribute to a body of knowledge or to make action-oriented decisions about educational options. Missing from this perspective was any conception of "design research." The assumption at that time seemed to be that educational innovations came from developers, engineers, or practitioners and the primary practical use of research was only in producing evidence of whether educational innovations are effective or not.

PASTEUR'S QUADRANT

The late Donald Stokes argued in *Pasteur's Quadrant* (1997) that Bush's compact was conceptually flawed because the scientific enterprise is not a one-dimensional progression from fundamental research to useful outcomes. Rather, it can be represented by a two-dimensional graph, with utility providing one axis and the fundamental/applied continuum providing the other. Stokes used the example of Louis Pasteur to argue that research can be both fundamental and useful. There are four quadrants to Stokes' graph, but it is Pasteur's quadrant that can lead to broadened support for basic research. Stokes provided an intellectual framework for a new relationship between science and society, one that recognizes utility as a driving force for science, rather than its automatic consequence. Figure 1 exhibits this dual dichotomy as a four-fold table and consider its cells or quadrants:

Figure 1

Research is inspired by:	Considerations of use?	
Quest for fundamental understanding?	NO	YES
YES	Pure basic research (Bohr)	Use-inspired basic research (Pasteur)
NO		Pure applied Research (Edison)

The upper left-hand cell or quadrant includes basic research that is solely guided by the quest for understanding without thought of practical use. It might be called 'Bohr's quadrant' in view of how clearly Niels Bohr's quest of atomic structure was a

> pure voyage of discovery, however much his ideas later remade the world. The lower right-hand cell might be called 'Edison's quadrant' in view of how strictly this brilliant inventor kept his co-workers in Menlo Park, the first industrial research laboratory in America, from pursuing the deeper scientific implications of what they were discovering in their headlong rush toward commercially profitable electric lighting. And the upper right-hand cell, for basic research that is use-inspired, deserves to be known as 'Pasteur's quadrant.' (Stokes, 1994)

Stokes noted that:

> The lower left-hand quadrant of this figure is not empty, and the fact that it is not helps make the point that our four-fold table is not simply a more elegant form of the basic-applied spectrum. Although some Washington pathologies belong here, such as policy research that is mounted to block action, most of the studies that belong in this quadrant are of highly particular phenomena, without any explanatory purpose or practical use in view. The bird-watchers who are grateful for the systematic studies of the markings and incidence of bird species that went into Peterson's Guide to the Birds of North America might want to call this 'Peterson's quadrant,' although this is too limited an example to warrant the name. (ibid.)

Implicitly reminiscent of John Dewey's appeal for constant interaction between theory and practice in science and education, Stokes' perspective offers new conceptions for education research. His elucidation of new relationships between use and understanding have provided a significant milieu for thinking and planning (Lagemann, 2000). Now, associated with the Stoke's conception of "Pasteur's quadrant," we have the two conceptions of scientific research and evidence-based practice. Scientific research situates itself directly in "Pasteur's quadrant." Evidence-based practice provides a possible link between the best research available and education practice.

Scientific Research and Evidence-Based Practice

Why are "scientific research" and "evidence-based practice" now receiving so much federal emphasis. The fundamental reason for the great interest is the belief that "scientific research" and "evidence-based practice" could serve as powerful agents for improvement, if not fundamental reform, of education. And the reason for deep concerns about the improvement of education is simply that developed

nations around the world are now in what Peter Drucker calls "The Knowledge Age."[1] Drucker says:

> Education will become the center of the knowledge society, and the school its key institution. What knowledge must everybody have? What is "quality" in learning and teaching? These will of necessity become central concerns of the knowledge society, and central political issues. In fact, the acquisition and distribution of formal knowledge may come to occupy the place in the politics of the knowledge society which the acquisition and distribution of property and income have occupied in our politics over the two or three centuries that we have come to call the Age of Capitalism.
>
> In the knowledge society, clearly, more and more knowledge, and especially advanced knowledge, will be acquired well past the age of formal schooling and increasingly, perhaps, through educational processes that do not center on the traditional school. But at the same time, the performance of the schools and the basic values of the schools will be of increasing concern to society as a whole, rather than being considered professional matters that can safely be left to "educators."

Given this unprecedented focus on the performance of schools, education has become both a state and a national obsession. However, in the United States, education has been constitutionally solely a state responsibility. The Elementary and Secondary Education Act of 1965 (ESEA) marked the first major funding by the federal government for K–12 public education. The political logic of ESEA was that in order to break the cycle of poverty across the nation, federal funding was needed to provide "compensatory education" for impoverished children wherever they lived.

In the No Child Left Behind Act of 2001 (NCLB) compensatory education has once again been used as the political basis for vastly expanding the federal presence across the nation, but this time not to simply break the cycle of poverty, but to raise education standards, accountability and performance for all children throughout the nation's public elementary and secondary schools. The ESEA Act included specific titles creating new roles or enlarging existing roles for research and development (e.g., the R&D Centers and Regional Educational Laboratories) and for dissemination and innovation support. The NCLB legislation

---

[1] Peter Drucker, "The Age of Social Transformation," *The Atlantic Monthly*, November 1994. Available on the Internet at: http://www.theatlantic.com/politics/ecbig/soctrans.htm

is remarkably different from ESEA in the fact that references to "scientifically-based research" are found almost everywhere throughout the act. There is an irony in this. Generally, those in the U.S. Congress have held a low opinion of education research in terms of its productivity and impact. Frequently in discussions with congressional representatives, comparisons have been made to the productivity and practical impact of research in the fields of agriculture or medicine.

During the 1970s and early 1980s, the Cooperative Extension Service of the U.S. Department of Agriculture was studied by the Department of Education as a possible model for emulation in order to connect research to education practice. (On close examination the agricultural model was abandoned for a number of reasons.) Today, our comparison model has shifted to the field of medicine. Tracing the various reasons for this shift goes beyond the scope of this paper, but clearly the reading research program carried out over the past three decades at NICHD has been extremely influential.

SCIENTIFIC RESEARCH IN EDUCATION

Let us review "Scientific Research in Education," the study authored by the National Research Council's Committee on Scientific Principles in Education Research.[2] In 2000, Representative Michael Castle (R-DE) introduced a bill to reauthorize OERI that included definitions of "scientifically valid quantitative methods" and "scientifically valid qualitative methods." The relatively narrow interpretation of scientific research found in these definitions alarmed many researchers. The National Educational Research Policy and Priorities Board sponsored this NRC study. The committee was challenged to articulate a more inclusive conception of the principles of scientific research in education and to derive implications of their findings for the future of a federal education research agency. In the following sections we focus on outlining the committee's description of scientific principles.

Key assumptions of the committee (paraphrased):

1. There is no one method or process that unambiguously defines science.
2. Many scientific studies in education and other field will not pan out.
3. It is possible to describe the physical and social world so that multiple observers can agree on what they see.

[2] The NRC report can be ordered or read online at: http://www.nap.edu/catalog/10236.html

4. The committee does not hold a simplistic notion that scientific quality alone will improve the use of such research in school improvement. Scientific quality and rigor are necessary, but not sufficient.
5. Scientific research in education is a form of scholarship that can uniquely contribute to understanding and improving education, especially when integrated with other approaches to studying human endeavors. Education is influenced by human ideals, ideologies, and judgements of value, and these things need to be subjected to rigorous — scientific and otherwise — examination. (NRC, 2002, pp. 25–26)

The committee asserts that:

> At its core, scientific inquiry is the same in all fields. Scientific research, whether in education, physics, anthropology, molecular biology, or economics, is a continual process of rigorous reasoning, supported by a dynamic interplay among methods, theories, and findings ... The accumulation of scientific knowledge over time is circuitous and indirect ... The scientific enterprise depends on a healthy community of researchers and is guided by a set of principles ... We conclude that six guiding principles underlie all scientific research, including education research. (p. 2.)

Stated tersely these six principles are: [3]

1. Pose significant questions that can be answered empirically.
2. Link research to relevant theory.
3. Use methods that permit direct investigation of the question.
4. Provide a coherent and explicit chain of reasoning from evidence to theory and back again.
5. Replicate and generalize across studies.
6. Disclose research to encourage professional scrutiny and critique.

The NRC report continues by observing that while all sciences share common principles, every field of study develops a specialization as the principles are applied. Education is multilayered and occurs within an intersection among several institutions, including schools, universities, communities and families. It is highly value-laden and involves a diverse array of people and political forces

[3] The reader might wish to compare these principles with those for Cronbach and Suppes' Disciplined Inquiry listed on page 2. The NRC report was dedicated to Lee J. Cronbach.

that shape its character. Physical, economic, social, cultural, historical, and other contextual factors as well as the complexity of teaching and learning all interact to influence educational processes and outcomes. Moreover, human volition and ethical considerations further complicate the field of study.

Design for the conduct of scientific research in education

The committee's discussion of research design is developed around three themes:

1. A variety of legitimate scientific approaches exist. Research designs evolve, as do the questions they address, the theories they inform, and the overall state of science.
2. Designs and methods must be carefully selected and implemented to best address the questions at hand. Some methods are better than others for particular purposes.
3. In order to generate a rich source of scientific knowledge in education that is refined and revised over time, different types of inquiries and methods are required.

To simplify the presentation, the committee organized the discussion of design around three interrelated types of questions (and provided elaboration for subtypes of designs within each of the three question types) as follows:

1. What is happening? (estimates of population characteristics, simple relationships, descriptions of localized educational settings)
2. Is there a systematic effect? (study of causal relationships when randomization is feasible, causal relationships when randomization is not feasible)
3. Why or how is it happening? (exploring relationships when theory is fairly well established, exploring relationships when theory is weak)

This discussion of research designs is moderately extensive and deserves careful study. Nevertheless, the details are not essential for the purposes of this paper. However the following observations of the committee are definitely worth noting.

Observations about the current state of education research

1. There are a number of areas in education practice and policy in which basic theoretical understanding is weak.
2. Many large-scale educational policies and programs are undertaken without an adequate evidentiary base to inform their development,

implementation, or refinement over time.... Systematic study is needed about the ways that programs are implemented in diverse educational settings. "We view implementation research — the genre of research that examines the ways that structural elements of school settings interact with efforts to improve instruction — as a critical, underfunded, and unappreciated form of education research. We also believe that understanding how to 'scale up' (Elmore, 1996) educational interventions that have promise in a small number of cases will depend critically on a deep understanding of how policies and practices are adopted and sustained (Rogers, 1995) in the complex U.S. education system."

3. "In all this work, more knowledge is needed about causal relationships. In estimating the effects of programs, we urge the expanded use of random assignment.... [W]e also urge that randomized field trials be supplemented with other methods, including in-depth qualitative approaches that can illuminate important nuances, identify potential counterhypotheses, and provide additional sources of evidence for supporting causal claims in complex educational settings."
4. "In sum, theory building and rigorous studies in implementation and interventions are two broad-based areas that we believe deserve attention." [pp. 125–126]

Although the NRC report deals primarily with examination of the character of scientific research in education, there is a chapter on design principles for fostering science in a federal education research agency (written in anticipation of the reauthorization of OERI). We shall not review that chapter's content.

Commentary and critiques of the NRC report

The committee's report succeeded in provoking substantial discussion in the education research community. The *Educational Researcher* published a theme issue on Scientific Research in Education (Volume 31, Number 8, November 2002).[4] Michael Feuer, Lisa Towne, and Richard Shavelson provide an introduction describing the NRC report and also reply. Critiques are provided by: James Pellegrino and Susan Goldman; David Berliner; Frederick Erickson and Kris Gutierrez; and Elizabeth St. Pierre. We briefly summarize each of these critiques.

> James Pellegrino and Susan Goldman in "Be Careful What You Wish For — You May Get It: Educational Research in the Spotlight" make two major

[4] The Educational Researcher theme issue can be accessed at: http://www.aera.net/pubs/er/toc/er3108.htm. The individual articles in this issue can be downloaded in PDF from this site.

points. The authors agree with Feuer, Towne and Shavelson that legislative mandates supporting particular methodologies would be misdirected. “It is bad policy to legislate scientific methods because issues of scientific method and quality are far too complex, contextualized, and nuanced.” Rather, legislation should mandate rigorous peer review of federally funded work and fund a peer review infrastructure within federal agencies. Their second point is that the educational research field has to work much harder to recognize and overcome the present state of fractionalization. Individual researchers and institutions of higher education must do more to model, support and reward the kind of multi-disciplinary, researcher-practitioner team approach found in the medical sciences. “[A]self-monitoring professional community is sorely needed to generate the body of high-quality research that contributes to understanding educational processes and outcomes and that makes a difference for the everyday lives of educators and students.”

David Berliner — “Educational Research: The Hardest Science of All.” Berliner asserts that educational science is unusually hard to do and that the government may not be serious about wanting evidence-based practices in education. Berliner notes that the “evidence-based practices” and “scientific research” mentioned so many times in NCLB legislation are code words for randomized experiments. This legislation confuses the methods of science with the goals of science. The NRC committee recognized the mistake. Berliner elaborates on his claim that education research is the “hardest-to-do-science” by providing examples regarding the power of context, the ubiquity of interactions, and the problem of “decade by findings” interactions [the short half-life of findings]. “Instead of putting its imprimatur on the one method of scientific inquiry to improve education, the government would do far better to build our community of scholars as recommended in the NRC report ... We should never lose sight of the fact that children and teachers in classrooms are conscious, sentient, and purposeful human beings, so no scientific explanation of human behavior can ever be complete.”

Frederick Erickson and Kris Gutierrez — “Culture, Rigor, and Science in Educational Research.” In this article the authors argue that both Feuer, Towne and Shavelson and the NRC report must be understood in the context of current federal discourse that focused narrowly on experimentally derived causal explanations of program effectiveness. The authors are concerned that the NRC committee, by accepting uncritically its charge to define scientific research produced a statement that could be read as endorsing “an evidence-based social engineering approach to educational improvement nationwide.” They express concern with ED’s Strategic Plan that holds up evidence-based medicine for emulation. We need “a more complicated and realistic

view of what actual scientists do and the varied and complex methods and perspectives they employ in their inquiry."

Elizabeth Adams St. Pierre — "Science Rejects Postmodernism." St. Pierre argues that despite claims to present an inclusive view of science, the NRC report narrowly defines science as positivism and methodology as primarily quantitative. The NRC definitions reject postmodernism and omit other theories "including queer, feminist, race, postcolonial, critical, and poststructural theories." St. Pierre uses postmodern analyses to illustrate the danger of the report's normalizing and totalizing discourse as an attempt to marginalize certain epistemologies and methodologies. She urges researchers to be on guard to keep educational research an open field in science that eschews conformity and methodological zealotry but encourages the proliferation of knowledge

It seems ironic that the committee's effort to broaden the definition of "scientific research" well beyond that in the Castle bill would be met with so many concerns within the education research community. However, *Scientific Research in Education* made it clear to many researchers that scientific principles were important. It was just that the traditional research principles espoused by the NRC committee seemed to rule out, or as Elizabeth St. Pierre states it, "marginalize," some epistemologies and methodologies. Other critics such as Erickson and Gutierrez perceive a larger problematic issue in the move toward "an evidence-based social engineering approach." Pellegrino and Goldman, and also Berliner, recognize that legislative mandates and definitions can pose dilemmas. As Pellegrino and Goldman observe, "It is bad policy to legislate scientific methods because issues of scientific method and quality are far too complex, contextualized, and nuanced." With this comment in mind, let us examine details of current legislative definitions.

SOME LEGISLATIVE LANGUAGE

We looked at some NSF definitions. Now let's examine some legislative definitions.[5]

From the NCLB Act:

Scientifically-based research "... means research that involves the application of rigorous, systematic, and objective procedures

[5] We may frequently ignore legislative definitions. However, the following definitions deserve special attention because they spell out congressional intentions and expectations for education R&D work. The NRC authors note that *Scientific Research in Education* did have an influence in broadening the definitions found in the Education Sciences Reform Act.

> to obtain reliable and valid knowledge relevant to education activities and programs." (NCLB Act)

And from the Education Sciences Reform Act of 2002 (P.L. 107–279), first note that this legislation preserved the traditional distinction of basic and applied research, as well as the definitions for development and dissemination (Sec. 102. Definitions).

> The term 'applied research' means research —
>
> (A) to gain knowledge or understanding necessary for determining the means by which a recognized and specific need may be met; and
> (B) that is specifically directed to the advancement of practice in the field of education.
>
> The term 'basic research' means research —
>
> (A) to gain fundamental knowledge or understanding of phenomena and observable facts, without specific application toward processes or products; and
> (B) for the advancement of knowledge in the field of education.

But note the emphasis on "scientifically valid" in the following two definitions.

> The term 'development' means the systematic use of knowledge or understanding gained from the findings of scientifically valid research and the shaping of that knowledge or understanding into products or processes that can be applied and evaluated and may prove useful in areas such as the preparation of materials and new methods of instruction and practices in teaching, that lead to the improvement of the academic skills of students, and that are replicable in different educational settings.
>
> The term 'dissemination' means the communication and transfer of the results of scientifically valid research, statistics, and evaluations, in forms that are understandable, easily accessible, and usable, or adaptable for use in, the improvement of educational practice by teachers, administrators, librarians, other practitioners, researchers, parents, policymakers, and the public, through technical assistance, publications, electronic transfer, and other means.

Moreover, the legislation introduces some new standards definitions including:

> SCIENTIFICALLY-BASED RESEARCH STANDARDS —
>
> (A) The term 'scientifically-based research standards' means research standards that —
>
> (i) apply rigorous, systematic, and objective methodology to obtain reliable and valid knowledge relevant to education activities and programs; and
>
> (ii) present findings and make claims that are appropriate to and supported by the methods that have been employed.
>
> (B) The term includes, appropriate to the research being conducted —
>
> (i) employing systematic, empirical methods that draw on observation or experiment;
>
> (ii) involving data analyses that are adequate to support the general findings;
>
> (iii) relying on measurements or observational methods that provide reliable data;
>
> (iv) making claims of causal relationships only in random assignment experiments or other designs (to the extent such designs substantially eliminate plausible competing explanations for the obtained results);
>
> (v) ensuring that studies and methods are presented in sufficient detail and clarity to allow for replication or, at a minimum, to offer the opportunity to build systematically on the findings of the research;
>
> (vi) obtaining acceptance by a peer-reviewed journal or approval by a panel of independent experts through a comparably rigorous, objective, and scientific review; and
>
> (vii) using research designs and methods appropriate to the research question posed.
>
> SCIENTIFICALLY VALID EDUCATION EVALUATION —
>
> The term 'scientifically valid education evaluation' means an evaluation that —
>
> (A) adheres to the highest possible standards of quality with respect to research design and statistical analysis;
>
> (B) provides an adequate description of the programs evaluated and, to the extent possible, examines the relationship between program implementation and program impacts;
>
> (C) provides an analysis of the results achieved by the program with respect to its projected effects;

(D) employs experimental designs using random assignment, when feasible, and other research methodologies that allow for the strongest possible causal inferences when random assignment is not feasible; and
(E) may study program implementation through a combination of scientifically valid and reliable methods.

SCIENTIFICALLY VALID RESEARCH —

The term 'scientifically valid research' includes applied research, basic research, and field-initiated research in which the rationale, design, and interpretation are soundly developed in accordance with scientifically-based research standards. *(P.L. 107–279, Sec. 102. Definitions)*

Evidence-Based Practice

Evidence-based practice (EBP) is an approach that argues that policy and practice should be justified in term of sound evidence about their likely effects. Currently, the term evidence-based practice is used with two different meanings. One meaning is associated with "best-practice." The other meaning is associated with practitioner decision-making (Mullen, 2002).

EBP as Best-Practice

An evidence-based practice is any practice that has been established as being effective through scientific research that conforms to some set of explicit criteria. For example, the selection criteria might include the following: (1) the practice has been standardized through manuals, guidelines, or certified training in the practice, (2) the practice has been evaluated through controlled research designs, (3) objective measures were employed that demonstrated valued outcomes, and (4) these outcomes have been replicated by different research teams.

EBP as Practitioner Decision-Making

This conception is attributed to David Sackett (Sackett, et al., 1996) who described evidence-based medicine as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients." Subsequently, Sackett and colleagues (2000) added that EBM is the "integration of best research evidence with clinical expertise and patient values." Thus, this conception places the emphasis not on a best-practice, but rather on a

decision-making process in which judgements are made on a case-by-case basis using the best-evidence available.

There are at least three challenges implied in either approach, namely the development of evidence-based policies, the development of evidence-based practice, and the promotion of a "culture of evidence" that would pervade the work of researchers and practitioners. However, how this is accomplished will depend on the whether the emphasis is placed on best-practices or on practitioner decision-making using best-evidence. If the emphasis is placed on "what works" or best practices, findings about those practices are disseminated for use by practitioners through agency directives, guidelines, manuals, accreditation requirements, or regulations. (Cf. NCLB legislation.) However, if a "culture of evidence" is to be developed successfully among practitioners, then much greater attention must be given to educating practitioners in critical decision-making in actual practice contexts.

In the United States, evidence-based practice movements have been underway for approximately three decades in the fields of medicine, employment, and welfare policy. And in the field of agriculture, it stretches back for almost a century. In all of these fields, rigorously controlled experiments and randomized field trials are considered the "gold standard" for evaluating the effectiveness of treatments or interventions. Because the U.S. Department of Education seems to be trying to emulate practices in medicine, we shall examine that movement.

EVIDENCE-BASED PRACTICE IN MEDICINE

In November 1992, an article in the *Journal of the American Medical Association (JAMA)* titled, "Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine" a new paradigm for medical practice was asserted, evidence-based medicine (EBM). The foundations for this paradigm shift were based on developments in clinical medical research over the previous 30 years. In 1960, the randomized clinical trial (RCT) was rare. By 1990 it was accepted that virtually no drug could enter clinical practice without demonstration of its efficacy in clinical trials. And RCTs were increasingly being applied to surgical therapies, diagnostic tests and other clinical practices. Moreover, meta-analyses were gaining acceptance as a method of summarizing the results of many similar RCTs. A new philosophy followed from these methodological advances. During the 1980s, there were many articles instructing clinicians how to access, evaluate, and interpret the expanding medical research literature.

The traditional clinical paradigm put a high value on established scientific authority and adherence to standard approaches. Answers to clinical questions were most frequently sought from direct contact with local experts or by

reference to the writings of national or international experts. However, the new EBM paradigm calls for additional clinical skills, including:

> ... precisely defining a patient problem, and what information is required to solve the problem; conducting an efficient search of the literature; selecting the best of the relevant studies, and applying rules of evidence to assess their validity; being able to present to colleagues in a succinct fashion the content of the article, and its strengths and weaknesses; extracting the clinical message, and applying it to the patient problem. (*JAMA*, 1992)

Thus, evidence-based medicine involves skills of problem definition, searching, evaluating, and applying the findings of original medical research literature. Between 1992 and 2000, *JAMA* published two dozen Users Guides to the medical literature. These guides provide clinicians with strategies and tools to interpret and integrate evidence from published research in their patient care. However, in acknowledgement of practical time demands of clinicians, systematic summaries, clinical guidelines, decision analyses, clinical pathways, and other EBM tools were increasingly employed to summarize original research findings succinctly and authoritatively. As *JAMA* authors developed these guides, their understanding of EBM evolved. By 2000, in a guide discussing principles of applying the guides to patient care, the guide concluded:

> The Users' Guides to the medical literature provide clinicians with the tools to distinguish stronger from weaker evidence, stronger from weaker syntheses, and stronger from weaker recommendations for moving from evidence to action. Much of the Guides are devoted to helping clinicians understand study results and enumerate the benefits, side effects, toxicity, inconvenience and costs of treatment options, both for patients in general and for individual patients under their care. A clear understanding of the principles underlying evidence-based practice will aid clinicians in applying Users' Guides to facilitate their patient care. Foremost among these principles are that value judgements underlie every clinical decision, that clinicians should seek evidence from as high in the appropriate hierarchy [of evidence] as possible, and that every clinical decision demands attention to the particular circumstances of the patient. Clinicians facile in the use of the Guides will complete a review of the evidence regarding the clinical problem with the best estimate of benefits and risks of management options and a good sense of the strength of inference concerning those benefits and risks. This leaves clinicians in an excellent position for the final — and still inadequately explored

– steps in providing evidence-based care, the consideration of the individual patient's circumstances and values. (Guyatt et al. 2000)

THE MEDICAL MODEL

The first medical article based on a randomized trial was published in 1948. And until 1962, the randomized trial was rare and controversial in medicine. However, in 1962, Congress passed the Kefauver-Harris amendments to the Food, Drug and Cosmetic Act, which included a requirement that before a new drug could be put on the market, there had to be "adequate and well-controlled investigations showing that the drug was effective." At the time both the pharmaceutical industry and the AMA opposed this requirement. Nevertheless, the Food and Drug Administration interpreted the requirement to require two independent randomized trials before FDA would grant a license for a new drug to go on the market.

The medical knowledge base

By 1966 the number of randomized trials in medicine rose to 100 per year. The pharmaceutical industry and the medical profession began to embrace the randomized trial as the "gold standard" for evidence-based medicine. This FDA requirement along with decisions at the National Institutes of Health to begin funding large-scale randomized trials massively increased the employment of randomized trials in medicine. By 1995 the number of randomized trials exceeded 10,000 per year. Today, well over a quarter million randomized trials are recorded by the Cochrane Collaboration and other medical research databases. Now we should recall that the pharmaceutical and medical device industries are multibillion dollar competitive enterprises that are forced to spend millions of dollars in randomized clinical trials in order to bring a single new drug or medical device to market. Funding for research at NIH exceeds $27 billion annually.[6]

The clinical decision setting

Although there are exceptions, the typical setting for EBM is a medical clinic where an attending physician examines an individual patient.[7] The key processes are diagnosis, prescription (treatment), and perhaps development of a prognosis

[6] http://www.nih.gov/news/budgetfy2004/fy2004presidentsbudget.pdf. However, note that most of the NIH budget is not used to fund clinical trials. Rather, it is used to create the knowledge needed to get to clinical trials.

[7] The unit of treatment may be individual patients, groups of patients, hospitals, HMOs, public health areas or other aggregations. However, it is usually the individual patient.

for the disease. Each of these processes can be informed by medical research findings. Evidence-based medicine involves conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients (Sackett, et al., 1996). Thus the practice of EBM calls for integrating individual clinical experience, usually based on years of medical training and clinical experience, with the best available evidence from systematic research. The key assumptions are (1) that there is a classification of diseases, (2) that a precise, correct diagnosis can be made, (3) that there are one or more treatment options for the diagnosed disease, and (4) that there is research-based evidence about the outcomes for each treatment.

There are two fundamental principles of EBM:

- >> Assessment of validity posits a hierarchy of evidence to guide clinical decision making
- >> The decision maker must always trade off benefits and risks, inconveniences and costs, and in doing so consider the patient's values and preferences

The hierarchy of evidence commonly recognized in EBM is as follows (Guyatt et al., 2000):

1. N of 1 randomized trial (patient serves as own control in alternating, double blind, pairs of target and alternative treatments — possible for only some kinds of medical treatments)
2. Systematic reviews of randomized control trials (meta-analyses and syntheses)
3. Single randomized control trial
4. Systematic review of observational studies addressing patient-important outcomes
5. Single observational studies addressing patient-outcomes
6. Physiologic studies
7. Unsystematic clinical observations

The second fundamental EBM principle stresses that "evidence is never enough." The clinician must always consider the magnitude of potential benefits and risks associated with alternative strategies, the likelihood that research-based outcomes (both benefits and risks) for trial populations can be particularized to a specific patient, and how to elicit and incorporate societal and individual patient values and preferences. Addressing these issues constitute an enormous challenge for EBM.

By far the biggest obstacle to the practice of EBM is the limited time a practicing clinician can afford to expend searching for the best available evidence. Fortunately, many resources have been developed to assist clinicians. Guyatt et al. (2000) provide a "4S" classification of information sources, namely:

- Primary Studies (preprocessing involves selecting only those studies that are highly relevant and have study designs that minimize bias, thus permitting a high strength of inference)
- Summaries (systematic reviews that provide clinicians with an overview of all the evidence addressing a focused clinical question)
- Synopses (synopses of individual studies or of systematic reviews that encapsulate the key methodological details and results required to apply the evidence to individual patients)
- Systems (summaries that link a number of synopses related to the care of a particular patient problem. These may take the form of practice guidelines, clinical pathways, or evidence-based textbook summaries of a clinical area)

Increasingly, easy electronic access to all these levels of evidence-based resources have become available, along with strategies and tools to interpret and integrate evidence from published research in patient care. However, this approach to providing clinicians with research has meant winnowing out virtually all medical research except the abbreviated compiled results of randomized or controlled clinical trials.

EBE — EVIDENCE-BASED EDUCATION

Given the emphasis on "scientifically-based research" found in NCLB, the U.S. Department of Education has taken action to support a movement for evidence-based practice in education. It recently funded the What Works Clearinghouse (WWC) to evaluate research and collaborated with the Coalition for Evidence-Based Policy in the preparation of the report, "Bringing Evidence-Driven Progress to Education: A Recommended Strategy for the U.S. Department of Education." While the report is that of the Coalition,[8] it is the product of extensive discussions with Department officials and staff.

The report notes that the NCLB legislation with its central principle that federal funds should support educational activities backed by "scientifically-based research"

[8] The Advisory Board for the Coalition included: Robert Boruch, Jonathan Crane, David Ellwood, Judith Gueron, Ron Haskins, Robert Hoyt, Blair Hull, David Kessler, Diane Ravitch, Laurie Robinson, Isabel Sawhill, Martin Seligman, Robert Slavin, Robert Solow, and Nicholas Zill.

offers an opportunity to bring rapid, evidence-driven progress to U.S. elementary and secondary education. To address the opportunity, the Coalition for Evidence-Based Policy undertook a collaborative initiative with the Education Department to explore how the Department can most effectively use its new authority to advance. The Coalition's report calls for a major, Department-wide effort to:

1. Fund studies that randomly assign students to treatment and control groups, in order to establish what works in educating children, and
2. Provide strong incentive for the widespread use of educational practices proven effective in such randomized controlled trials.

The rationale is that this strategy would be key to reversing decades of stagnation in education and sparking rapid, evidence-driven progress.

The report urges the Department to make a concerted effort to support randomized controlled trials, support a knowledge base of these proven interventions, and spur their wide-spread use in order to fundamentally improve the effectiveness of American education. The key recommendations in the report are as follows.

Recommendation to create the infrastructure within the Department for this proposed strategy:

- The Department should identify "High Priority Areas" in which there is a critical need to (i) build the knowledge base of proven interventions or (ii) provide incentives for their widespread use.
- To fund the randomized trials and other recommendations below, the Department should deploy the following programs and funding sources to the maximum extent practicable:
  - (i) Office of Education Research and Improvement [now Institute of Education Sciences] funds for research and dissemination;
  - (ii) "National activities" funds (for evaluation, demonstration, dissemination, and technical assistance) that the Department is authorized in law to carry out in many grant programs and areas of policy;
  - (iii) A small percentage allocation from Elementary and Secondary Education Act programs, similar to that implemented by the Justice Department with informal approval of Congressional appropriators.

Recommendation to build the knowledge base of effective, replicable interventions in High Priority Areas:

- The Department should focus its discretionary funds for research and evaluation on randomized trials to identify such interventions.

>> The Department's grant program should give applicants major incentive to use their discretionary funds to carry out such randomized trials. This would enable the Department to leverage a much larger pool of resources to carry out this strategy. Such incentives would include: (i) additional funding for the applicant from the funding sources listed above, (ii) in discretionary grant programs, significant additional points in the proposal evaluation process; (iii) waiving of certain statutory and/or regulatory requirements (a strategy used effectively by the Department of Health and Human Services to get states to test welfare programs in randomized trials); and (iv) positive recognition and publicity.

Recommendation to provide strong incentives for the widespread use of proven interventions:

>> The Department's grant program should require applicants to provide a plan for widespread implementation of research-proven interventions, with quantifiable goals. This would apply to both formula and discretionary grant programs (discretionary programs would make this plan an important factor in the proposal evaluation process). Importantly, each applicant would be responsible for choosing which interventions, backed by randomized trials, to include in its plan.

In High Priority Areas, the Department would require an independent evaluation, after grant award, of whether the applicant meets the goals of its plan. The Department would annually issue a high-profile report summarizing the results of these evaluations, including the progress of each state agency and other major grantee in implementing research-proven interventions.

The Coalition's report concludes that the above recommendations can all be implemented within the Department of Education's existing statutory authority. It further notes that implementation of the recommendations "will require a major, sustained commitment from the Department, involving the ongoing coordination and strategic deployment of programs and resources." We need to take these recommendations seriously[9] considering that the report was developed in collaboration with Department of Education officials and that Secretary Page introduced the report at a major policy forum with senior officials from ED, HHS, Labor, and Justice; OMB; congressional committees; and education advocacy groups.

[9] The ED press release for the Secretary of Education's remarks opens with the following statement: "The U.S. Department of Education has come one step closer to ensuring that teaching and learning in the nation's classrooms are based on solid, empirical educational practices. Under a joint effort with the department, the Coalition for Evidence-Based Policy today [11/18/02] issued a report calling for a major, department-wide effort to fund studies that randomly assign students to treatment and control groups, to establish what works in educating America's children."

We also note that ED's Performance Measures for Objective 4.1, Quality and Rigor of Department-funded Research, includes:

> **Use of Randomized Experimental Designs**
>
> **Projects.** Of new research and evaluation projects funded by the Department that address causal questions, the percentage that employ randomized experimental designs.* Performance Targets: FY 02 Base Line +10 PP FY 03 Base Line +25 PP [PP = Percentage Points]
>
> **Publications.** Of new research and evaluation publications funded by the Department that address causal questions, the percentage that describe studies that employ randomized experimental designs.* Performance Targets: FY 02 Base Line +10 PP FY 03 Base Line +25 PP

EVIDENCE-BASED EDUCATION AT THE INSTITUTE OF EDUCATION SCIENCES

As of June 2003, we found PowerPoint presentations and transcripts of several presentations. Below and on the following pages is copied the content of a presentation on EBE by Grover Whitehurst. [http://www.ed.gov/offices/OESE/SASA/eb/evidencebased.pdf]

1. Evidence-Based Education (EBE)
   Grover J. (Russ) Whitehurst, Assistant Secretary, Educational Research and Improvement, [Now Institute of Education Sciences] United States Department of Education]

2. Three Stories
   - The university president
     + Evidence isn't relevant
   - The vendors
     + What constitutes evidence isn't clear
   - Teaching
     + Evidence isn't available

3. What is EBE?
   The integration of professional wisdom with the best available empirical evidence in making decisions about how to deliver instruction

4. What is professional wisdom?
   - The judgment that individuals acquire through experience
   - Consensus views
   - Increased professional wisdom is reflected in numerous ways, including the effective identification and incorporation of local circumstances into instruction

5. What is empirical evidence?
   - Scientifically-based research from fields such as psychology, sociology, economics, and neuroscience, and especially from research in educational settings
   - Empirical data on performance used to compare, evaluate, and monitor progress

* These would include all research and evaluation studies initiated by any office within the Department, but would exclude collections of statistics.

6. Evidence-based Education
   - Professional Wisdom
     + Individual Experience
     + Consensus
   - Empirical Evidence
     + Scientifically-Based Research
     + Empirical Information

7. Why are both needed?
   - Without professional wisdom education cannot
     - adapt to local circumstances
     - operate intelligently in the many areas in which research evidence is absent or incomplete.
   - Without empirical evidence education cannot
     - resolve competing approaches
     - generate cumulative knowledge
     - avoid fad, fancy, and personal bias

8. Medicine and Ag as Models
   - A little history
   - Evidence-based medicine
   - Examples
     - The Illinois Library
     - The FTC and diet pills
     - Hormone Replacement Therapy [HRT] Study

9. The HRT Study
   - Sample: 27,000+ Women, aged 50–79.
   - Research Design: Women randomly assigned to receive either hormone therapy or a placebo; Data collected for 8–12 years.
   - Hypothesis: HRT will reduce heart disease and fractures without increasing breast cancer

10. The HRT Study [continued]
    Table Description:
    This chart reflects disease rates for women on estrogen plus progestin or placebo based on the results of a Hormone Replacement Therapy Study designed to show the number of cases of heart attacks, strokes, breast cancer, blood clots, colorectal cancer, hip fractures, endometrial cancer and deaths per year in 10,000 women. Overall statistics reflect an increase in heart attacks, strokes, and breast cancer for women on estrogen plus progestin for heart ailments.

11. Social Policy and ED examples
    - Nurse-home visitation
    - DARE
    - High quality preschool
    - National Reading Panel report

12. Policy Requirements
    - Difference in the mix of professional judgment, scientific research, and objective measures that justifies imposition of requirements contrasted with identification as good practice
    - Reading research vs. math research as example

13. Scientifically Based Research
    "...means research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs"
    (No Child Left Behind Act of 2001)

14. Scientifically Based Research
    - Quality
      - To what degree does the design and analysis and logical inference support the claims and conclusions?
    - Relevance
      - To what degree are the variables and circumstances similar across the research and the settings in which the research is to be applied?

15. Quality: Levels of evidence
    All evidence is NOT created equal
    1. Randomized trial (true experiment)
    2. Comparison groups (quasi-experiment)
    3. Pre-Post comparison
    4. Correlational studies
    5. Case studies
    6. Anecdotes

16. Randomized Trials: The gold standard
    - Claim about the effects of an educational intervention on outcomes
    - Two or more conditions that differ in levels of exposure to the educational intervention
    - Random assignment to conditions
    - Tests for differences in outcomes

17. Why is randomization critical?
    - Assures that the participants being compared have the same characteristics across the conditions
    - Rules of chance mean that the smart, motivated, experienced, etc. have the same probability of being in condition 1 as in condition 2

18. Why is randomization critical? (Continued)
    Without randomization, simple associations such as between Internet use and science grades have many different interpretations
    Table Description: "Average science scores by students' reports on use of the Internet at home" [This chart indicates an increase in science scores by students' reports on use of the Internet at home for grades 4, 8 and 12.]

19. Relevance
    - Does the study involve a similar intervention and outcome to those of interest?
    - Were the participants and settings representative of those of interest?

20. Evidence will not make the decision
    - Be skeptical
    - Consider other ways of achieving goal
    - Consider consequences and local circumstances
    - Consult with experts who understand evidence before making costly decisions (This is different from consulting authorities who may know the subject area but not rules of evidence)

21. EBE – Where are we? Description:
    This pie chart graphic shows a circle indicating external evidence is only a [very] small slice of the pie compared to professional wisdom as related to evidence based education.

22. Education Lags Behind
    Chart Description: This chart indicates the total number of articles about randomized field trials in other areas of social science research (criminology, social policy and psychology) has steadily grown over the last 40 years; however, the number related to education research has trailed behind. [By approximately 1996, the cumulative number of articles about definite and possible randomized field trials in criminology is approaching 6,000; the numbers in social policy and psychology exceed 2,000; while the number for education is less than 1,000.]

23. What ED will do
    The What Works Clearinghouse (w-w-c.org)
    - interventions linked to evidentiary support
    - systematic reviews
    - standards for providers of evaluations, and list of evaluators who have agreed to follow those standards

24. What ED will do
    The National Center for Education Evaluation
    - Well designed, timely, & nonpartisan evaluations of ED's own programs
      + Funding streams
      + Specific interventions
    - Funding for development and evaluation of interventions in the field
    - Feedback into discretionary grant programs

25. What ED will do
    - Internal review of ED's own products
    - Build capacity in the field
    - Professional training
    - Workshops for major decision makers
    - Systematic and long-term research programs to fill gaps

26. Goals
    - ED will provide the tools, information, research, and training to support the development of evidence-based education
    - The practice of evidence-based education will become routine
    - Education across the nation will be continuously improved
    - Wide variation in performance across schools and classrooms will be eliminated

We leave to the reader the challenge to review the above presentation and arrive at a conclusion about the balance in perspectives between viewing EBE as Best Practice or EBE as Practitioner Decision-Making. (See page 14.)

The What Works Clearinghouse (WWC)

Side 23 mentions the What Works Clearinghouse. Copied below is the description currently found at the Institute of Education Sciences' web page.

(See http://www.ed.gov/offices/IES/NCEE/wwc.html)

> **What Works Clearinghouse (WWC)**
>
> The What Works Clearinghouse (WWC) has been established by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, and the public with a central, independent, and trusted source of scientific evidence of what works in education. It is administered by the Department through a contract to a joint venture of the American Institutes for Research and the Campbell Collaboration.
>
> Educators, policymakers, and the public need a central, trusted, and independent source of evidence about what really works in education. To meet this need, the WWC develops standards for reviewing and synthesizing educational research, selects topic areas for review, and conducts systematic reviews of existing research. The WWC will provide its findings in accessible, user-friendly, searchable on-line databases that will include the following:
>
> - reviews of potentially replicable interventions (i.e., programs, products, and practices) that are intended to enhance student outcomes;
> - information about the evaluation studies on which intervention reviews have been based;
> - scientifically rigorous reviews of test instruments used to assess educational effectiveness; and
> - a registry of evaluators (individuals and organizations) willing to conduct quality evaluations of education interventions.
>
> The WWC develops standards for scientific evidence on educational effectiveness and conducts and publishes systematic reviews of existing research. To ensure independence and high quality, the work of the Clearinghouse is advised by a Technical Advisory Group (TAG) which is composed of the nation's leading experts in research design, program evaluation, and research synthesis. The TAG advises on the standards for the

research syntheses, monitors and informs the methodological aspects of the research, and reviews the evidence reports.

The WWC seeks broad participation from all those interested in improving the nature and the role of evidence in education and is committed to ensuring that its products and services meet user needs. The WWC collaborates with a large network of producers and consumers of research evidence to ensure broad input into WWC plans and activities.

For more information on the What Works Clearinghouse: www.w-w-c.org

[This IES/NCEE web page was last modified—May 21, 2003].

EBE IMPLICATIONS

Where are we in education versus medicine? The quick answer is, "At least 40 years behind medicine." Currently, it is obvious that major emphasis at the Institute of Education Sciences and elsewhere in the Department of Education is on significantly increasing the number of randomized control trials (RCTs). However, recall that the Kefauver-Harris amendments to the Food, Drug and Cosmetic Act were passed in 1962 (see page 17). The NCLB Act was passed in 2001 and the Education Sciences Reform Act was passed in 2002. Moreover, there is nothing in education that parallels the economic motivations of the pharmaceutical and medical device industries to fund thousands of RCTs in order to bring their products to market.

The Campbell Collaboration's registry claims to currently contain over 10,000 entries on randomized trials and *what seem to be randomized trials*[10] in psychology, education, and criminology (see Appendix, page 41). However, the chart depicted in slide 18 (see page 22) indicates that perhaps less than 1,000 of the registry entries are actually in the field of education. The WWC has just begun to conduct reviews of potentially replicable interventions and reviews of test instruments. Then again, the EPPI Centre in England has been at work developing systematic reviews for almost a decade (see Appendix, pages 41–45). In the United States and England there are strong government-led movements to establish evidence-based practices in many social fields, including education. Nevertheless, it may take a decade or more to begin to approach (qualitatively but not quantitatively) where medicine was in the early 1990s.

[10] Inspection of a sampling of the Campbell Collaboration abstracts of studies in the field of education suggests that some of the studies included in the registry may not in fact be based on randomized treatments. We may suspect is that over time a much larger proportion of studies added to the registry will be true randomized studies.

Currently, the emphasis at the U.S. Department of Education and the Institute of Education Sciences seems to be on Evidence-Based Education "as best-practice" or "what works." This is the primary thrust of the recommendations made by the Coalition for Evidence-Based Policy (see pages 19–22). In Russ Whitehurst's presentation (see especially slides 3–7, 12, 19–20, and 25–26) one can discern the possibility of some eventual support for EBE "as practitioner decision-making." Yet, it is not at all clear when or to what degree this latter interpretation of EBE will evolve. The primary emphasis seems to be on "what works."

This emphasis is understandable given the top-down, mandatory perspective on affecting educational practitioner behavior that seems to pervade the NCLB legislation. Moreover, there is the practical need to significantly increase the number of RCTs in education if we are to develop a trustworthy basis for identifying and supporting "proven" practices. However, if we can learn anything from experience in the development of Evidence-Based Medicine, or indeed from Evidence-Based Practices in other fields (Davies, Nutley & Smith, 2000), much more thought and effort will need to be given to considering how to effectively change the behavior, and indeed the culture, of practitioners in the field. Before EBE can truly be successful, it must succeed in promotion of a "culture of evidence" that pervades practitioner decision-making at all levels of education throughout the nation.

That requires a substantial investment in EBE education, training, mentoring, and continuing education. Tens of thousands of RCTs, hundreds of systematic reviews, and many types of summary systems (practice guidelines, practice pathways, evidence-based journals and textbooks) will only serve to create the beginning of an authoritative and hopefully accessible knowledge base. However, an authoritative knowledge base alone can do little to change practice among key decision-makers, policy-makers, and opinion leaders throughout the field of education. That will require creating an authentic "culture of evidence." Is this an immense task? Of course it is. But that is exactly what EBE is fundamentally about: changing the practice of education.

To focus primarily on simply building a large, very high quality "what works" knowledge base would be another instance of taking the "linear view," not unlike that of "first research, then development, then dissemination, and finally application." Several decades ago, we learned that effective development work needed to authentically involve all types of intended users from the very beginning and at every stage through out the effort. When we failed to do that, we ran the substantial risk that the product failed to meet user needs or preferences. It was difficult to market and difficult to bring to scale because there was not sufficient consumer interest or acceptance.

In the case of EBE we are not thinking in terms of a product or a simple program or practice, but rather in terms of a very large professional and cultural change

movement. But the same requirement exists. Indeed, even after several decades in the field of medicine, EBM is still hard at work attempting to penetrate more deeply into clinical practice, and with only moderate degrees of success. But perhaps we especially need to pay attention to the slogan in EBM that "Evidence is never enough!" (See page 14.) Moreover there are significant problems with the concept of "evidence" in the field of education.

Problems with evidence

The concept of "evidence" is problematic.[11] Evidence is not value-free. Try to get a roomful of teachers, politicians or researchers to agree on the meaning of even something as apparently simple as "effective practice." Opponents of the "evidence-based" approach cite the value-laden nature of all "evidence." There are no universal solutions or quick fixes. Education is so complex that subtle differences in contexts can make large differences in the effects or outcomes of changes in policy or practice. This contextual effect makes it unlikely that simple, universal strategies will produce the improvements intended everywhere. A more useful kind of evidence that is required is detailed information that would identify exactly which features of the context are important in affecting various desired outcomes.

Evidence is often incomplete or equivocal. Too often politician take the stance that they need to act, or at least be seen to be acting, despite the existence of any clear evidence about what action might be appropriate. In many education policy areas we simply do not really know enough to support a clear decision. However, decisions must be made. If the concept of using evidence as a basis for policy and practice is to prevail, then we need to avoid making exaggerated claims based on limited knowledge. Where evidence is equivocal, we need to explore the nature of the conflicting evidence and conduct additional experiments to try to resolve the ambiguity. Evidence can be very complex. When statistical analyses are involved, especially in complex experimental designs, where there are interactions among variables, the evidence can be difficult to interpret simply.

Moreover, as David Berliner points out in his critique of *Scientific Research in Education*, it is not simply the ubiquity of interactions, but also the preponderant influence of contexts.[12]

---

[11] The following two paragraphs are drawn from a discussion of problems with "evidence" in education found at http://cem.dur.ac.uk/ebeuk/problems.htm

[12] The Tennessee STAR study (on class size reduction) is recognized as one of the classical randomized studies in education. Yet one only needs to examine the history of the implementation of class size reduction in California to appreciate the hazards in ignoring the importance of contexts in generalizing from the findings of a study to application in another context.

There is a substantial body of material in the field of Evidence-Based Medicine, including several users guides, that attempts to make clinicians aware of the risks in generalizing (particularizing) the findings of medical research RCTs to particular patients in specific settings.

A far more daunting problem exists in education because the character of the "treatment," the causal factors and interactions, the social complexity of the settings, and the dynamic nature of the educational process are almost always far more complex than treating an individual patient for a specific disease in a clinic or hospital. Medicine is hardly simple, but it differs from education along at least the following dimensions:

- Decision process
- Character of the "treatment"
- Unit of treatment (school or classroom versus individual patient)
- Contexts
- Causal factors and interactions
- Static character of a drug or surgical intervention versus dynamic human activity systems
- Social complexity of the classroom versus physiological complexity of single patient

EBE is in some ways similar to EBM **in Public Health**. In a recent evaluation of a Norwegian program to promote evidence-based public health practice (Forsetlund, et al., 2003, 200–201), the evaluators observe:

> In Norway, tasks within general public health are often about providing background information and advice for local health decisions, decisions that are taken in a political setting. The reasons the physicians in this study gave for not referring to research in policy documents express the existing norms of their social system, as well as their own norms. . . Interestingly, the physicians themselves seem to work in the same way as do policy makers and take the similar issues into consideration in their decision making. The analysis of interview data indicated a strong resemblance to how Weiss found that policy makers seek out and use research information. Like the policy makers, the physicians did "relatively little search for evidence or analysis. People tend to make do with what they already know — or at least know about" (Weiss, 1986, 276).

> There has been some skepticism about the promotion of evidence-based policy simply as an extension of evidence-based medicine, that is, policy making is qualitatively different.

By citing the above quotation, we do not wish to imply that we should view the decision-making tasks of educational practitioners as similar to those of public health physicians. Rather, we are simply suggesting that decision making among many educational practitioners could also be "qualitatively different" given the bullet list presented immediately above.

Problems with applying research evidence to policy

The study just cited goes on to reference an editorial by Black (2001). In this editorial Black notes that a useful distinction has been made between practice policies (use of resources by practitioners), service policies (resource allocation, pattern of services), and governance policies (organizational and finance structures).

Concerning practice policies Black observes that the relation between research evidence and clinical practice has been thoroughly examined by practitioners of EBM. The linear, rationalist model holds up well, although it shows signs of strain in two areas. First, policymakers differ in their interpretation of evidence. Second, there is a lack of generalisability as EBM moves away from drugs to manual interventions, e.g., surgery, because decisions depend on features of the specific patient (obesity, anatomy, quality of tissues), the particular surgeon, and various external factors (equipment available, competence of assistants). [Is there any similarity here to the teacher in the classroom?]

Regarding service policies, Black states that the relation between research evidence and service policies is generally weak. He lists (and discusses) six main reasons:

1. Policymakers have goals other than clinical effectiveness (social, financial, strategic development of service, terms and conditions of employees, electoral)
2. Research evidence is dismissed as irrelevant (from different sector or specialty, practice depends on tacit knowledge, not applicable locally)
3. Lack of consensus about research evidence (complexity of evidence, scientific controversy, different interpretations)
4. Other types of competing evidence (personal experience, local information, eminent colleagues' opinions, medicolegal reports)
5. Social environment not conducive to policy change
6. Poor quality of knowledge purveyors

Regarding governance policies, Black notes that the direct influence of research on governance policies has been negligible. He states: "Clearly, research has only a limited role because governance policies are driven by ideology, value judgments, financial stringency, economic theory, political expediency, and intellectual fashion. It would be naïve and unrealistic to expect research to provide evidence to clinch arguments about governance policies."

Black draws several conclusions from his discussion of practice, service and governance policies.

> Firstly, research has little direct influence on service and governance policy if we adopt criteria set and accepted by researchers. Secondly, the relation between research and policy depends on the arena and, thus, the policymakers. Research evidence is more influential in central policy than local policy, where policymaking is marked by negotiation and uncertainty. Thirdly, the use of research depends on the degree of consensus on the policy goal. It is used if it supports the consensus and is used selectively if there is a lack of consensus. Fourthly, many researchers are politically naïve. They have a poor understanding of how policy is made and have unrealistic expectations about what research can achieve. And fifthly, policymaking is not an event but is "ethereal, diffuse, haphazard and somewhat volatile." The consequences of failing to understand this are clear: "So long as researchers presume that research findings must be brought to bear upon a single event, a discrete act of decision making, they will be missing those circumstances and processes where, in fact, research can be useful." In other words, we need a better model to underpin the relation.

Black then describes the enlightenment model (Weiss, 1977), where research is seen as but one of several knowledge sources. In the Weiss model research can provide policymakers with new ways of conceptualizing the world, formulating policy issues, mapping the decision making terrain, or challenging conventional assumptions. Black goes on to note that:

> During the 1980s and 1990s this view was extended to a more interactive model based on a close dialogue between researchers and policymakers in which knowledge is considered to be inherently contestable. ... It is necessary, therefore, to consider which arguments are likely to be useful or gratifying to which policymakers. Researchers have to accept that their work may be ignored because policymakers have to take the full complexity of any situation into account. They need to recognize that

other legitimate influences on policy (social, electoral, ethical, cultural, and economic) must be accommodated and that research is most likely to influence policymakers through an extended process of communication.

Please note that Professor Black's editorial is titled, "Evidence based policy: proceed with care." He writes from experiences in the field of health care. But it seems that his observations about the various forms of policy (practice, services, and governance) contain cautionary lessons that may also be applicable to the field of education.

Some general observations about implementation of EBM and EBE

What might we add to the above critiques about problems with evidence and applying it to policy settings? We should note that currently EBE seems to be relying primarily on a top-down, "push" strategy. And it seem to be following a highly linear, rational approach. Such approaches have almost always grossly underestimated the obstacles and barriers posed by cultural, behavioral, and organizational constraints on widespread adoption. Although EBM has a several decades start on EBE, both movements face major dissemination, marketing, implementation and scale-up problems. Impediments to implementation scale-up exist in medicine, especially in lack of clinician time to keep up with the research; however, medicine has many positive supports and forces that favor eventual success of the EBM movement.

On the other hand, there are far fewer supports in education. There is a presumption that education will eventually become a "performance-driven" field. But one may ask how soon or to what extent will that be true. Our conclusion is that for EBE to become successful, very significant and extensive amounts of "intermediation" must be provided (e.g., EBE products and services). And a substantial amount of practitioner education will be needed (education not only in "EBE practice" but more fundamentally in professional practice decision making).[13]

[13] A reviewer noted that professional wisdom varies enormously among education practitioners. The author agrees. And that is why a substantial amount of practitioner education will be needed.

Implications for Field-based Work

EVIDENCE-BASED PRACTICE

Although this paper has focused on the importance placed on EBE at the U.S. Department of Education and the Institute of Education Sciences, the evidence-based practice movement is a general thrust in federal government. The Office of Management and Budget has made it clear to various federal departments and agencies that it wants to see much more compelling evidence of performance and accomplishment reported in the annual GEPRA reports. Moreover, there will be increasing demands for "scientific evidence" to be provided to support the development and implementation of programs at many agencies. Consequently, contractors can expect to see increasing emphasis placed on the quality of the "evidence" they may be asked to provide concerning much of the work they do for any federal agency, not just for ED or IES.

This emphasis is likely to affect work across a wide spectrum, including research, development, evaluation, dissemination, technical assistance, and training. Initially, much of the concern will be centered on identifying "what works," providing rigorous evidence that "it works," and perhaps specifying for whom, and under what circumstances does a product, program, or practice work. Having satisfied those demands, and given more sophistication (on the part of the sponsor and the contractor), we can expect to encounter additional concerns about "why does it work."

This is what we can expect as long as the federal emphasis is focused primarily on the "push" or supply side of evidence-based policy and practice. However, any applied research and service agency should anticipate that eventually there will evolve a much greater concern at federal agencies with promoting evidence-based policy and practices in the field. This will be a new and much more comprehensive conception of "dissemination" and "implementation" — one that will be concerned with affecting policy making and practitioner decision making so that they are truly evidence-based. In the final two sub-sections of this paper we shall return to this challenge, but first we make a few observations about "scientific research."

THEORY-BASED PRACTICE (SCIENTIFIC RESEARCH)

The above heading telegraphs our message. The focus on "what works" is simply not enough in education or related social fields. The interventions themselves, the application contexts, and the interactions between interventions and contexts are usually too complex to safely generalize from even carefully controlled randomized experiments. We also need to know why it works, where it works, and under what conditions does it work. Now this is precisely what the NRC

committee observed about the current state of education research (see page 8). We should especially note the statement: "In sum, theory building and rigorous studies in implementation and interventions are two broad-based areas that we believe deserve attention." (Shavelson & Towne, 2000, p. 126)

Let's review those six principles set forth in Scientific Research in Education (see page 7). They are:

1. Pose significant questions that can be answered empirically.
2. Link research to relevant theory.
3. Use methods that permit direct investigation of the question.
4. Provide a coherent and explicit chain of reasoning from evidence to theory and back again.
5. Replicate and generalize across studies.
6. Disclose research to encourage professional scrutiny and critique.

Note that principles 2 and 4 both involve theory. One might ask why did the committee place so much emphasis on theory? There are a number of reasons, but perhaps the most compelling is that social research in general, and especially education research, has been criticized for failing to provide a coherent, cumulative knowledge base. Much of this problem can be traced to shoddy work in following all six principles, but especially to failures to link the research to relevant theory and failures in connecting research findings to that theory.

This is tough work even when the research is dealing with moderately "fundamental" or "basic" research topics. It become extraordinarily difficult when we are working in highly applied educational or social practice areas. In such settings, if there is theory, the work is often multidisciplinary. It may be related to several different forms of theory. And findings could be linked back to several theories. Although, all too often, there is little or no theory that explicitly guides the work or that the work contributes to. In other words, specifically Donald Stokes' words, the work is in "Edison's Quadrant." And to be realistic, that is where much of our educational research, development, and technical assistance work is located.

The push for more rigorous "scientific research" will compel us to find a way to move more of our work into "Pasteur's quadrant." That will not be easy. As Matthew B. Miles often said, "We have many theories of change, but few theories of changing." And we can look to his work in organizational development and in school reform for ideas about how to externalize tacit practitioner knowledge and create "thick" principle-guided theories of action for changing practice. Moreover, we can look to the movement in theory-driven evaluation as a way to construct useful theories of action and then test and refine them through evaluation feedback (Chen, 1990; Weiss, 1997).

EBE SUPPORT PRODUCTS AND SERVICES

What EBE Products and Services are likely to be needed? We might look at the product and services that already exist in medicine. Listed below are some of the more significant EBM products and services:

- Syntheses and synopses
- Systematic reviews (See Appendix, pages 41 and 44.)
- Meta-analyses
- Systematic syntheses (Precise quantification of all benefits and risks)
- Integrative articles
- Users' guides
- Practice guidelines (consensus panels)
- Decision analysis aids
- Clinical pathways (educational practice pathways)
- Evidence-based practice information systems & services
- Evidence-based journals and textbooks
- Evidence-based practice tutorials & training (web-based and traditional) (See Appendix, page 48.)
- Evidence-based practice courses

Some of these types of products and services are beginning to appear in education. See for example the work of the EPPI Centre (Appendix pages 41–44). Over time we can assume that, as the EBE movement is confronted with its own need for "evidence" that it is having positive effects on policy makers and practitioners in the field, there will be greater awareness of the need for many more of these kinds of products and services.

EBE ASSISTANCE TO PRACTITIONERS AND POLICYMAKERS

State and local policy makers, administrators, and other education practitioners are already confronting the many requirements posed by NCLB legislation regarding adoption of programs and practices that are "scientifically-based." And, as Professor Black observes:

> ... the relation between research and policy depends on the arena and, thus, the policymakers. Research evidence is more influential in central policy than local policy, where policymaking is marked by negotiation and uncertainty.... [And], the use of research depends on

> the degree of consensus on the policy goal. It is used if it supports the consensus and is used selectively if there is a lack of consensus.

Although these observations are based on experience in the fields of medicine and health, they appear to also hold true for the field of education. One size does not fit all. The diversity of states, districts, schools and communities found throughout the nation guarantees that state and local decision making will be "marked by negotiation and uncertainty." Moreover, when we are dealing with the education of children, various local values and preferences must be weighed against what research evidence may dictate. Local policy and decision making will be problematic especially when the research evidence is based on different circumstances, contexts, types of students, and perhaps employs outcome measures reflecting different values and preferences than those dominant in local communities. Under these circumstances policy makers and education practitioners are likely to appreciate assistance in interpreting research evidence in ways that honor the spirit of the NCLB law but also recognize local conditions and preferences.

On this point we should refer back to the second fundamental principle of EBM (see page 17):

>> *The decision maker must always trade off benefits and risks, inconveniences and costs, and in doing so consider the patient's values and preferences*

In discussing this principle Guyatt et al.(2000) state:

> Thus, knowing the tools of evidence-based practice is necessary but not sufficient for delivering the highest quality of patient care. In addition to clinical expertise, the clinician requires compassion, sensitive listening skills, and broad perspectives from the humanities and social sciences. These attributes allow understanding of patient's illnesses in the context of their experience, personalities, and cultures.

And those responsible for delivering the highest quality of education surely need similar attributes for understanding the students placed in their care. This will be the challenge for policymakers and educators everywhere as they strive to integrate professional wisdom with the best available empirical evidence in making decisions about education policy and practice. The challenge for applied research, development, and service agencies will be to provide assistance in developing and interpreting research evidence in ways that promote effective integration with professional wisdom.

References

Black, N. (2001). Evidence based policy: proceed with care. *British Medical Journal* (bmj) 2001:323:275–279 (4 August). http://bmj.com/cgi/reprint/323/7307/275.pdf

Bush, V. (1945). *Science The Endless Frontier: A Report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development.* Washington: DC: United States Government Printing Office. http://www.nsf.gov/od/lpa/nsf50/vbush1945.htm

Chen, H. T. (1990). *Theory-driven evaluation.* Newbury Park, CA: Sage Publications.

Coalition for Evidence-Based Policy (2002a). *Bringing Evidence-Driven Progress to Education: A Recommended Strategy for the U.S. Department of Education.* Report of the Coalition for Evidence-Based Policy. Washington, DC: Author. Report is available for download at: http://www.excelgov.org/usermedia/images/uploads/PDFs/coalitionFinRpt.pdf

Coalition for Evidence-Based Policy (2002b). *Rigorous Evidence: The Key to Progress in Education? Lessons from Medicine, Welfare, and Other Fields* (Transcript of the policy forum proceedings). Washington, DC: Author. Transcript is available for download in MS Word at: http://excelgov.org/usermedia/images/uploads/PDFs/Final_Online_Version_-_November_18_Forum_Transcript.doc

Cronbach, L. J. & Suppes, P. (Eds.) (1969). *Research for tomorrow's schools: Disciplined inquiry for education.* Report of the Committee on Educational Research of the National Academy of Education. New York: Macmillan.

Davies, H. T. O., Nutley, S. M. & Smith, P.C . (Eds.) (2000). *What Works? Evidence-based policy and practice in public services.* Bristol, UK: The Policy Press.

Elmore, R.F. (1996) Getting to scale with good educational practice. *Harvard Educational Review*; 66(1), 1–26.

Feuer, M & Towne, L. (2002) *The Logic and the Basic Principles of Scientific Based Research*. http://www.ed.gov/offices/OESE/esea/research/feuer-towne-paper.html

Forsetlund, L., Talseth, K.O., Bradley, P., Nordheim, L. & Bjørndal, A (2003). Many a slip between cup and lip: Process evaluation of a program to promote and support evidence-based public health practice. *Evaluation Research*, Vol. 27 No. 2, April 2003, 179–209.

Guyatt, G. H., Haynes, B., Jaeschke, R. Z., Cook, D. J., Green, L, Naylor, C. D., Wilson, M. C. & Richardson, W. S. (2000) EBM: Principles of Applying Users' Guides to Patient Care. *Journal of the American Medical Association*, 284 (10) 1290–1296. http://www.cche.net/usersguides/applying.asp

JAMA (1992). Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. *Journal of the American Medical Association*, (1992 Nov. 4; 268(17): 2420–5). http://www.cche.net/usersguides/ebm.asp

Lagemann, E. C. (2000). *The Spencer Foundation Annual Report, April 1, 2000 — March 31, 2001, Report of the President*. Chicago, IL. http://www.spencer.org/publications/annual_reports/ar_2001.pdf

Mullen, E. J. (2002). *Evidence-Based Knowledge: Designs for Enhancing Practitioner Use of Research Findings (a bottom up approach)*. Presentation at 4th International Conference on Evaluation for Practice, University of Tampere, Tampere, Finland. http://www.uta.fi/laitokset/sospol/eval2002/EvidenceF2002.PDF

NSF (1998). National Science Foundation, Division of Science Resources Studies, *National Patterns of R&D Resources: 1998*, NSF 99–335, by Steven Payson (Arlington, VA 1999). http://www.nsf.gov/sbe/srs/nsf99335/htmstart.htm

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York: Free Press.

Sackett, D. L., Rosenberg. W., Gray, J. M., Haynes, R. B. & Richardson, W. S. (1996). Evidence-Based Medicine: What it is and what it isn't. *British Medical Journal* (312: 71–72). http://bmj.com/cgi/content/full/312/7023/71

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W, & Haynes, R. B. (2000). *Evidence-based Medicine: How to Practise and Teach EBM* (Second Edition). Edinburgh: Churchill Livingstone.

Shavelson, R. J. & Towne, L. (Eds) (2002). *Scientific Research in Education*. (Prepared by Committee on Scientific Principles for Education Research, National Research Council.) Washington, DC: National Academy Press. (The report can be read online and hard copies are available for sale at http://www.nap.edu/catalog/10236.html) (The report is summarized in Feuer & Town (2002) above.)

Stokes, D. E. (1994). The Impaired Dialog Between Science and Government and What Might Be Done About It. From Albert H. Teich, A.H., Nelson, S.D. &McEnaney, C., (Eds). *Science and Technology Yearbook, 1994*, Washington: DC: American Association for the Advancement of Science, pp. 123–145. http://www.wws.princeton.edu/programs/stpp.articles/dialog.html

Stokes, D. E. (1997) *Pasteur's Quadrant: Basic Science and Technological Innovation*, Washington, DC: Brookings Institution Press. Available at Brookings Press Books Online: http://brookings.nap.edu/books/0815781776/html/R1.html

Weiss, C. H. (1979). Research for policy's sake: the enlightenment function of social research. *Policy Analysis*, 3:531–547.

Weiss, C. H. (1986). The circuitry of enlightenment. *Knowledge: Creation, diffusion, utilization*. 8:274–81.

Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review*, 21:501–24.

Appendix: Evidence-Based Practice Web Resources

GENERAL RESOURCES

EvidenceNetwork
http://www.evidencenetwork.org/home.asp

The Focus Point for Evidence Based Policy and Practice Research in the UK

What is EvidenceNetwork?

- >> A starting point for accessing social science research publications relevant to policy and practice
- >> Open to users in the research community, the voluntary sector, local and central government, public agencies and commercial organisations
- >> Providing search tools and a referral framework to enable users to pursue their enquiries
- >> A forum for debate and discussion of issues and problems in relation to evidence-based policy

(See History) http://www.evidencenetwork.org/history.asp

In 1999 the Economic and Social Research Council, the UK's largest funding agency for research and postgraduate training in social and economic issues, decided that a major initiative was needed to bring social science research much nearer the decision making process. The Council's Research Resources Board launched the Evidence Based Policy and Practice (EBPP) Initiative with overall funding of £3 million over three years for a national co-ordinating Centre and a network of research Nodes.

Following applications from a number of universities and research institutes, Queen Mary, University of London was selected as the national Centre and began work in December 2000. Seven institutions were successful in their applications to become Nodes, and an eighth was added in late 2002. Together, the Centre and Nodes make up EvidenceNetwork.

The Centre's objectives are to:

- >> provide co-ordination and support for the Nodes
- >> foster the exchange of research based evidence between policy researchers and practitioners, and increase their mutual understanding
- >> accelerate the development of methods for appraising and summarising

the results of research relevant to policy and practice across social science disciplines

- improve the quality of research, policy development and practice

The eight Nodes work in a range of social science disciplines, covering policy and practice issues in the fields of public health, economics (two Nodes), children, ethnic health, neighbourhoods, social care and research utilisation. Their objectives are to:

- focus on broadly based policy issues in the subject field
- have a flexible configuration, enabling change and variation over time in response to changing needs and priorities
- ensure that the Centre has access to all available expertise, data and information relating to their domains
- have an active role in EvidenceNetwork through participation in joint projects with each other, and operational interaction with the Centre

Coalition for Evidence-Based Policy

Information on the Coalition for Evidence-Based Policy and its activities can be found at: http://www.excelgov.org/performance/evidence/execsumm.htm

Useful Links (Evidence-Based Practice)
http://homepages.ihug.co.nz/~catkins/ebis/Resources/useful_links.htm

EDUCATION

Campbell Collaboration
http://www.campbellcollaboration.org

The international Campbell Collaboration (C2) is a non-profit organization that aims to help people make well-informed decisions about the effects of interventions in the social, behavioral and educational arenas. C2's objectives are to prepare, maintain and disseminate systematic reviews of studies of interventions. We acquire and promote access to information about trials of interventions. C2 builds summaries and electronic brochures of reviews and reports of trials for policy makers, practitioners, researchers and the public.

The Campbell Collaboration Library includes the:

- Register of Campbell Systematic Reviews of Studies of interventions in the Social, Behavioural and Education arenas

>> C2 Social, Psychological, Education, and Criminological Trials Registry (C2-SPECTR). http://128.91.199.101 Currently contains over 10,000 entries on randomized trials and what seem to be randomized trials. Unique in the world, C2-SPECTR is composed of abstracts on completed randomized experiments and on planned experiments. It is updated continuously. Its contents are part of the ingredients for The Collaboration's systematic reviews and the reviewers augment the contents.

>> The C2 Titles and Protocol Registries contain all approved Titles and Protocols (plans) for each systematic review proposed by review teams.

EPPI Centre (Evidence for Policy and Practice Information) http://eppi.ioe.ac.uk/EPPIWeb/home.aspx

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London. The EPPI-Centre was established in 1993 to address the need for a systematic approach to the organisation and review of evidence-based work on social interventions. The work and publications of the Centre engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Research Evidence in Education Library [at EPPI Centre] http://eppi.ioe.ac.uk/reel/

The Research Evidence in Education Library (REEL) is the home site of the Centre for Evidence-Informed Policy and Practice in Education, commissioned by the Department for Education and Skills, England. The Centre's vision is to be a centralised resource for people wishing to undertake systematic reviews of research in education and those wishing to use reviews to inform policy and practice. This part of the website has two distinct functions: to publish systematic reviews in education which have been written by members of review groups; and to provide tools and databases for those wishing to undertake systematic reviews in this field.

Evidence-Informed Education http://eppi.ioe.ac.uk/reel/home.aspx?page=/reel/about_evidence.htm

The current interest in policy and research circles in systematic reviews and evidence-informed education is part of a general move in the UK and elsewhere towards basing policy and professional practice on sound evidence of effectiveness. In the health sector, for instance, it has become clear that much of what health care professionals do is not derived from reliable evidence, and that sometimes what professionals believe in with all the best intentions may not only

be ineffective but sometimes actually harmful. This has led to systems being set up to ensure that professionals and policy-makers have constantly updated access to the findings of good quality research. The most notable of these is the Cochrane Collaboration, which aims to make syntheses of evidence on the effectiveness of healthcare interventions accessible to practitioners, policymakers and users.

Similar developments, being taken forward by a sister organisation to the Cochrane Collaboration, the Campbell Collaboration, are now taking place in relation to social policy, social welfare, criminal justice and other areas of 'social' intervention. There is much that researchers in education and users of educational research can learn from work in these other areas, although some of the challenges of research synthesis in education are particular to that setting.

Systematic Reviews

The education stream of work in the EPPI-Centre is focused on systematic reviews. These involve identifying research reports and reviewing them in an explicit and standard way so as to produce new and accessible syntheses of the evidence.

Systematic reviews in education can collate a range of research types so as to investigate: what works, and what doesn't: how things work or why they don't; or current practices, trends, needs or promising areas for development.

More information about systematic reviews:
http://eppi.ioe.ac.uk/reel/home.aspx?page=/reel/about_reviews.htm

What is a systematic review?

A systematic review is a piece of research. Like any piece of research, it uses research methods that aim to make it produce valid and reliable results. For example, systematic reviews include efforts to find as much as possible of the research which addresses the review's research question. This is important if the review's conclusions are not to be over-influenced by studies which are simply the easiest to find (usually published research, showing the benefit of interventions). Another example of the methodological approach of a systematic review is the use of a set of explicit statements, called inclusion criteria, to assess each study found to see if it actually does address a review's research question. As is the case for any good research, the methods for a systematic review are made explicit in a 'protocol' before it starts. This helps to reduce bias in the review process, for example, by ensuring that reviewers' procedures are not overly influenced by the results of studies they find. If changes are needed to the protocol as the review progresses these needed to be noted in the review's final report and the rationale for making changes made clear.

A systematic review is also explicit in reporting its methods so that these can be appraised. For example, the methods used to find studies (database searches, searches of specialist bibliographies, hand-searching of likely journals, attempts to track down unpublished research) will be reported in some detail. This allows readers to decide for themselves whether the reviewers have looked carefully enough to be able to say they have identified as many as possible of the studies that could help answer the review's research question. It is now standard practice for reports of systematic reviews to have clearly defined methods and results sections.

An important characteristic of a systematic review is that it includes a synthesis of its results, which in this case are results from previous research. As a very important part of the synthesis process, systematic reviewers assess the quality of the studies they have found. They can then use this assessment to assign different weights to study findings. Poor quality studies are sometimes downgraded in importance or excluded from the review. The ultimate effect of this is that research can influence a review's conclusions only when that research is sound.

The synthesis is usually presented in the form of structured narrative, summary tables or a statistical combination (meta-analysis). This synthesis is then used to formulate conclusions and recommendations. The aim is to make the links between the detail of the studies found and the reviewers' conclusions clear.

[This information is taken from the Review Group Manual, which is available online as an 82-page PDF working document, dated February 2001: http://eppi.ioe.ac.uk/EPPIWebContent/downloads/RG_manual_version_1_1.pdf]

[From the Review Group Manual]

> The current interest in policy and research circles in systematic reviews and evidence-informed education is part of a general move in the UK and elsewhere towards basing policy and professional practice on sound evidence of effectiveness. In the health sector, for instance, it has become clear that much of what health care professionals do is not derived from reliable evidence, and that sometimes what professionals believe in with all the best intentions may not only be ineffective but sometimes actually harmful. This has led to systems being set up to ensure that professionals and policymakers have constantly updated access to the findings of good quality research, most notably by the Cochrane Collaboration, which aims to make syntheses of evidence on the effectiveness of healthcare interventions accessible to practitioners and policy makers. Similar developments are now taking place in relation to social policy, social welfare, criminal justice and other

areas of 'social' intervention. There is much that researchers in education and users of educational research can learn from work in these other areas.

Aims of the Initiative

The Evidence Informed Policy and Practice in Education Initiative is based on the following vision of evidence-informed policy and practice in education:

- high quality, relevant reviews of research being accessible, foremost to *teachers*, but also to policy makers, students, parents, governors and others with an interest in education;
- collaboration that develops systematic review methodology for educational research and helps ensure the use of review findings;
- a research process that is open to scrutiny, criticism and development;
- a research process that values and takes steps to encourage participation, at all stages, by anyone with an interest in education.

One of the key products of the Initiative is an electronic library of quality assured systematic reviews of research in education that is to be made accessible to all stakeholders.

Reviews in Education

The EPPI-Centre supports people working in the field of education to write reviews. The early results of this programme of work are also available online at: http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/reel/reviews.htm

[The reviews available by May 22, 2003 are the following]

- A systematic review of the impact of summative assessment and tests on students' motivation for learning
- A systematic review of the impact of networked ICT on 5–16 year olds' literacy in English
- A systematic review of classroom strategies for reducing stereotypical gender constructions among girls and boys in mixed-sex UK primary schools
- A systematic review of the effectiveness of school-level actions for promoting participation by all students
- The impact of financial circumstances on engagement with post-16 learning: a systematic map of research

- The effect of travel modes on children's mental health, cognitive and social development; a systematic review (April 2001) [Soon to be available online]
- The impact of leadership and management on school achievement.

Fourteen groups have registered with the centre to date, with a further four due to start in 2003. The registered Groups' home pages is available at: http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/reel/reviewers.htm

[This URL contains links to *completed reviews,* proposals and protocols for the following:]

- Art and design, Assessment and learning research, Citizenship, Continuing professional development, Early years, English teaching, Gender and education, Inclusive education,
- Modern languages, Post-compulsory education, School and the community: transitions,
- School leadership, Science education, and Thinking skills.

Evidence Based Education UK
http://www.cem.dur.ac.uk/ebeuk/

Contains links to conferences, publications, and related links.

No Child Left Behind web site
http://www.NoChildLeftBehind.gov

Promoting Research and Evidence-Informed Practice (REIP)
http://www.tta.gov.uk/itt/providers/research/

The Teacher Training Agency (TTA) is active in supporting the Government in its drive to promote teaching as a research and evidence-informed profession as a means of improving teaching and learning and raising standards.

The TTA works with a range of schools, Local Education Authorities and Higher Education Institutions and others to increase teachers' interest in and opportunities to engage with research and evidence at a local and national level. It aims to influence both the supply and demand for research and evidence with and on behalf of teachers. The agency gathers teachers' perceptions of research, runs conferences to challenge their beliefs about what research can do for them and they can do for research, and funds School Based Research Consortia who are piloting innovative ways of involving teachers at all stages of the research process, from commenting on the relevance of research priorities and proposals

through to dissemination and interpretation of findings. It has also provided small grants to teachers to carry out increasingly rigorous case study research aimed at increasing the interest of their colleagues in research and has been influential in persuading other organisations to do the same and exploring new ways of making research relevant and accessible to teachers.

The TTA uses its limited research budget on work which concentrates on exemplifying new ways of working, promoting teachers' needs within the debate about education research and persuading others to do the same. In this context it has funded three large scale research projects on effective teachers of literacy and numeracy and effective use of ICT in the teaching of these subjects.

Research and Evidence-Informed Practice [at TTA]
http://www.tta.gov.uk/itt/providers/research/research_evidence.htm

The Teacher Training Agency is developing a network of teachers to investigate practical ways of using research evidence in the classroom in order to improve standards.

Improving Standards: Research and Evidence Based Practice (pdf:19Kb) is designed to encourage teachers at local and national levels to make good use of research and evidence. It describes briefly the different opportunities for teachers to be involved in using and doing research which have been established by the TTA, and invites enquiries from teachers who are interested in finding out more.

The paper *Deputy/Headteachers' Views on Accessing and Using Research and Evidence - results of a pilot survey* (word doc:108Kb) provides details of initial work undertaken to gather teachers' views on what research has to offer them. The majority of views collected to date have been from headteachers and deputy heads and the TTA is gathering further evidence to broaden the sample of teachers' views identified.

SOCIAL INTERVENTIONS

Evidence for Policy and Practice Information Co-ordinating Centre
http://eppi.ioe.ac.uk/EPPIWeb/home.aspx

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London. The EPPI-Centre was established in 1993 to address the need for a systematic approach

to the organisation and review of evidence-based work on social interventions. The work and publications of the Centre engage health and education policy

makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings. An increasing number of the publications are made available on this website.

SOCIAL WORK

Evidence Based Social Services
http://www.ex.ac.uk/cebss/

The Centre for Evidence-Based Social Services (CEBSS) is jointly funded by The Department of Health and a consortium of Social Services Departments in the South West of England with the main aim of ensuring that decisions taken at all levels in Social Services are informed by trends from good-quality research. CEBSS is based at the University of Exeter and is part of the Peninsula Medical School.

OUR MAIN AIMS

- To translate the results of existing research into service and practice development;
- To ensure research findings are available to Social Services Departments when reviewing and changing service delivery, and are fed into the review process;
- To collaborate with DipSW, Degree and PQ course providers to ensure that education and training in social work incorporates the knowledge available from existing research;
- To improve the general dissemination of research findings to local policy makers, managers, practitioners and service users and careers;
- To commission new research where significant gaps are identified;
- To work towards the inclusion of service-users and careers, as a particularly valuable source of information on service-effectiveness projects.

Additional aims of the project include to improve evidence-based skills in Social Services.

MEDICAL

The Centre for Evidence-Based Medicine, Oxford
http://www.cebm.net

A site produced by the Curriculum Evaluation Centre at the University of Durham. The site discusses similarities with Evidence Based Medicine and provides conference details and some tools such as "What is an effect size?" (This site provides some excellent links to other evidence based resources in medicine and healthcare. It includes links to resources for learning or doing EBM and includes a Toolbox with calculators.)

Cochrane Collaboration
http://www.cochrane.de/cc/default.html

Preparing, maintaining and promoting the accessibility of systematic reviews of the effects of health care interventions. Several databases are included in The Cochrane Library. One of them, The Cochrane Database of Systematic Reviews, contains Cochrane reviews and another, The Cochrane Controlled Trials Register, is a bibliographic database of controlled trials. The Database of Abstracts of Reviews of Effects (DARE) includes structured abstracts of systematic reviews which have been critically appraised by reviewers at the NHS Centre for Reviews and Dissemination in York and by other people, e.g. from the American College of Physicians' Journal Club and the journal Evidence-Based Medicine. The Cochrane Methodology Register is a bibliography of articles on the science of research synthesis. Also included in The Cochrane Library is a Reviewers'Handbook on the science of reviewing research; a Glossary of methodological terms and Cochrane jargon; and contact details for review groups and other groupings in the Cochrane Collaboration.

[Note access to Systematic Reviews and the Register is by subscription. However, there is free access to the approximately 1,600 Abstracts of Reviews of Effects. To review abstracts by Collaborative Review Groups (CRGs) use the following: http://www.cochrane.de/cc/cochrane/revabstr/mainindex.htm]

Cochrane Consumer Network
http://www.cochraneconsumer.com

The Cochrane Consumer Network's site contains a range of health care information, and information to help people understand health care research. It is also a resource for consumers and others who want to become involved in the Collaboration or other health research activities.

Evidence-Based Medicine Retrospective References
http://www.herts.ac.uk/lis/subjects/health/ebmrefs.htm

This is a selective, but substantial, listing of references on the theory and methodology of evidence-based medicine/healthcare from 1993 to 2002.

Note that 2003 references are on the main evidence-based medicine/healthcare page at http://www.herts.ac.uk/lis/subjects/health/ebm.htm#refs

WISDOM
http://www.wisdomnet.co.uk

WISDOM is a pilot project based at the University of Sheffield and funded by the National Health Service Executive to create an on-line environment, using the Internet to train primary care professionals in informatics. At the heart of the project is a discussion group: this web site supports the group and offers information resources and background to the project. Evidence Based Practice is one of the focus areas for the project (Practice because it targets all members of the Primary Health Care Team). Several tutorials, originally distributed by e-mail but now deposited in the project archive, cover aspects of evidence based practice.

WISDOM Tutorials to date cover:

- Evidence based practice 1: An Introduction to Evidence-Based Practice
- Evidence based practice 2: The basic stages in EBP and how to get started.
- Evidence based practice 3: A little bit about databases...
- Evidence based practice 4: Running a literature search
- Evidence based practice 5: How to Evaluate the Evidence
- Evidence based Practice 6: From Evidence to Practice
- Second series - EBP Seminar 1: Finding Useful Web Sites
- Second series - EBP Seminar 2: Critical Appraisal
- Second series - EBP Seminar 3: Relative & Absolute Risk Interpretation
- Second series - EBP Seminar 4: Clinical Governance : An Introduction

Netting the Evidence
http://www.shef.ac.uk/~scharr/ir/netting

Netting the Evidence is intended to facilitate evidence-based healthcare by providing support and access to helpful organizations and useful learning resources, such as an evidence-based virtual library, software and journals.

[Main subsections of this site include: Library, Searching, Appraising, Implementing,

Software (medical information and statistics calculators), Journals, Databases, Organisations, Search, and an A-Z Index]

[The Organizations link provides URLs for more than 50 evidence-based organizations.]

The Library (Core Library for Evidence Based Practice) is a virtual library with approximately 90 links to full text documents on all aspects of Evidence Based Practice. This is an unusually rich resource that may be useful to anyone interested in learning more about evidence-based practice in any field, not just nursing or medicine. Covers many topics such as how to read a paper, systematic reviews, meta-analysis, understanding controlled trials, getting research into practice, and much more.

Evidence Based Nursing — An e-Journal
http://ebn.bmjjournals.com

The general purpose of Evidence-Based Nursing is to select from the health related literature those articles reporting studies and reviews that warrant immediate attention by nurses attempting to keep pace with important advances in their profession. These articles are summarised in “value added” abstracts and commented on by clinical experts. The specific purposes of Evidence-Based Nursing are:

- To identify, using predefined criteria, the best quantitative and qualitative original and review articles on the meaning, cause, course, assessment, prevention, treatment, or economics of health problems managed by nurses and on quality assurance
- To summarise this literature in the form of ”structured abstracts” that describe the question, methods, results, and evidence-based conclusions of studies in a reproducible and accurate fashion

- To provide brief, highly expert comment on the context of each article, its methods, and clinical applications that its findings warrant
- To disseminate the summaries in a timely fashion to nurses.

Currently (May 2003) the following are free to non-subscribers:

- All tables of contents
- All abstracts
- Free sample issue
- Full text of "Editor's Choice" for the current issue
- Email alerts — get the table of contents and more, for each issue of this and other journals delivered to your mailbox (requires registering with HighWire Press)
- Search EBN Online or across 350+ HighWire journals
- Browse EBN Online archive of past issues
- Direct access to Medline
- Details of forthcoming events and announcements
- eLetters — lively correspondence about recent articles