Comment: Culture, Rigor, and Science in Educational Research
Frederick Erickson and Kris Gutierrez
EDUCATIONAL RESEARCHER 2002 31: 21
DOI: 10.3102/0013189X031008021

The online version of this article can be found at:
http://edr.sagepub.com/content/31/8/21

Published on behalf of

American Educational Research Association

and

http://www.sagepublications.com

Additional services and information for Educational Researcher can be found at:

Email Alerts: http://er.aera.net/alerts
Subscriptions: http://er.aera.net/subscriptions
Reprints: http://www.aera.net/reprints
Permissions: http://www.aera.net/permissions
Citations: http://edr.sagepub.com/content/31/8/21.refs.html

>> Version of Record - Nov 1, 2002

What is This?
In this article the authors argue that both the Feuer, Towne, and Shavelson article and the larger National Research Council (NRC) report on which it is based must be understood in the context of current federal discourse that focuses narrowly on experimentally derived causal explanations of educational program effectiveness. Although the authors concur with much of the Feuer et al. article and the NRC report, they are concerned that the NRC committee, by accepting uncritically its charge to define the scientific in educational research, produced a statement that risks being read as endorsing both the possibility and the desirability of taking an evidence-based social engineering approach to educational improvement nationwide. Finally, the authors review the consequences of not challenging the layperson’s “white coat” notion of science and replacing it with a more complicated and realistic view of what actual scientists do and the varied and complex methods and perspectives they employ in their inquiry.

There is much with which we agree in “Scientific Culture and Educational Research” (this issue) and even more in the National Research Council (NRC) report (2002) on which it is based. First, the article and the report state forthrightly that evidence-careful descriptive research done primarily by sustained firsthand observation and interviewing—sometimes called qualitative or case study or ethnography—can make valuable contributions to educational research, and that evidence-careful descriptive research falls within the range of methods in education that can be called scientific. Second, the article and the report make many reasonable recommendations for the organization of the Department of Education as a federal government sponsor of research. They note that such research sponsors should be partially insulated from political pressures, but that in a democracy that insulation can never be total nor should it be. We suggest two lines of addition and correction. The first concerns technical matters regarding qualitative research that were mentioned in passing in the article and report but which we believe deserved greater emphasis there. The second concerns the basic premise of the NRC committee’s charge.

One technical matter is that the article and the report both emphasize causal analysis by means of experiment in order to determine educational program effects, as part of the current federal and state system of accountability for effectiveness. But the article and the report do not make it clear that in the very enterprise of determining program effects, qualitative research is more than merely allowable; it is essential if causal analysis is to succeed. A logically and empirically prior question to “Did it work?” is “What was the ‘it’?”—“What was the ‘treatment’ as actually delivered?” Educational treatments are situated and dynamically interactive (see Cohen, Raudenbush, & Ball, 2002). They are locally constructed social ways of life involving continual monitoring and mutual adjustment among persons, not relatively replicable entities like chemical compounds or surgical procedures or hybrid seed corn or manufactured airplane wings. High fidelity implementation is rare in education—for reasons of local exigency—and despite the accountability pressures and the wishes of experimenters to avoid this major threat to internal validity, there are real-world limits on how “faithful” the implementation will be of even the most structured of instructional programs. Unless considerable proportions of a research budget, even in a large-scale formal experiment, are devoted to documenting the treatment as delivered on the ground, the causal inferences drawn from inspection of outcome data will remain unwarranted, and they are likely to be partially misleading half truths. “What was the treatment, specifically?” is a question best answered by qualitative research. It is expensive and absolutely necessary, especially if the push to scale up continues.

A second technical matter is that although the NRC report acknowledges that qualitative research can provide explanations for causal processes, it does not emphasize how crucial such insights can be for the enterprise of causal analysis: the Educational Researcher (ER) article is even less direct on this point. Light and Pillemer (1982), for example, discuss a study of day care centers in which quantitative results suggested that the most important influence on the children’s educational development was neither group size nor the staff-to-student ratio, which were the independent variables operationally defined and measured in the study. When observers went to look in the day care centers they found more than one adult in the larger groups of children, while in the smaller groups there was only one adult. In the larger groups the adults talked to one another, thus paying less attention to the children. In other words staff-to-staff schmoozing was the actual proximal independent variable, not the overall staff-to-student ratio or the size of instructional groups, in and of themselves. At issue was not the simple amount of a unitary “resource” such as group size or staff-to-student ratio as general program features but the specific interactive use of such potential resources within the conduct of instructional practice. Direct observation was necessary to detect and explain this. (We are indebted for this example to Joseph Maxwell, 1996; see also the discussion in Cohen et al., 2002.)

A second and more basic line of criticism concerns the premise of the charge to the NRC committee in the first place, which was to define what was scientific in educational research. Some
pronouncements on the scientific by the Department of Education and by members of Congress, such as Representative Michael Castle of Delaware, indicate that within the executive and legislative branches of the federal government a leap of faith has been taken toward belief in the unmixed blessings of hard science—causal analysis by means of experiment—as the only way to improve educational research. These beliefs are truly alarming in their naiveté and zeal.

Feuer et al. reference without critical comment the Department of Education’s Strategic Plan for 2002–2007, noting that it is available at the federal government publications website (www.ed.gov/pubs/stratplan2002-07/index.html). Every member of AERA should visit this website and read the language of Strategic Goal 4: “Transform education into an evidence-based field.”

The statement begins (p. 50), “Unlike medicine, agriculture, and industrial production, the field of education operates largely on the basis of ideology and professional consensus. As such, it is subject to fads and is incapable of the cumulative progress that follows from the application of the scientific method and from the systematic collection and use of objective information in policy making. We will change education to make it an evidence-based field.” Concerning Objective 4.1, “Raise the quality of research funded or conducted by the Department” (p. 51), the list of implementation strategies includes “improve the peer review process by articulating clear standards and by enlisting only those qualified scientists who have high levels of methodological and substantive expertise” and “create an editorial review board for all Department research publications to ensure that they meet the highest standards of scientific rigor before their publication.” Under Objective 4.2, “Increase the relevance of our research in order to meet the needs of our customers” (p. 52), an orienting comment says, “the Department will create and regularly update an online database of scientifically rigorous research on what works in education.”

Note that in the first quoted sentence “the scientific method” stands unqualified as the term in contrast to “ideology,” “professional consensus,” and “fads.” The next sentence says, “We will change education to make it an evidence-based field,” but it is clear that only one kind of evidence is to be given currency. Consider the effect in the next two quoted sentences concerning Objective 4.1 of substituting scholars or experts for “scientists” and scholarly rigor or rigorous research for “scientific rigor.” Consider the effects in the last quoted sentence about an on-line database of saying that it would contain high quality research on issues of current importance in education. Instead the database will show us “scientifically rigorous research on what works in education” without close and thorough examination of how and why those treatments “work.”

In addition, we have each witnessed officials of the U.S. Office of Educational Research and Improvement comparing current educational practices with those of medieval medicine (e.g., bleeding to cure fever) and characterizing the current state of knowledge in education as superstition. We are not postmodernists, but we must protest that this “high modernist” position constitutes a substitution of scientism, an idealization of science, for science itself. The prescription of a “scientific culture” as an effective remedy for the ills of educational research and of “hard science” causal studies of program effects as a remedy for defects in education practice must be treated very skeptically. Reluctantly, because we like so much of the ER article and the NRC report, we must argue that they cannot be read apart from the context of the broader federal discourse concerning the field of education. In that context, by not challenging the reigning optimism about hard science as a royal road to improvement, the article and committee report could be read as supporting a discourse of scientism or the appearance of rigor in educational research rather than its actual substance.

The problem is that the house is burning around us and the article and report never cry “Fire!” or “Danger!” Recall that the NRC committee’s charge was not to define standards for rigor, reason, and quality in educational research but to say what was scientific. And although the report and article tried to range broadly across an intellectual playing field that was profoundly tilted at the outset, they never questioned the fundamental premise of their charge. Indeed, the ER article goes farther down this slippery slope than does the committee report: As Feuer et al. say in so many words, “We therefore use terms like science, research, scholarship, and inquiry as essentially interchangeable in the specific context of discussing the norms and ideals of the educational research field” (p. 5). Furthermore, although the article and report mention Lakatos and Musgrove (1970) in passing, they do not mention at all the substantial line of work since the 1970s called social studies of science, which documents the ways in which actual scientists’ practice differs from that of its idealized characterizations (see, e.g., Latour & Woolgar, 1986; Lynch, 1993; Knorr-Cetina, 1999). From Lakatos on, many scholars have shown that real scientists in their daily work are anything but disinterested and canonically rational. In their daily practice they are passionate and argumentative, profoundly selective in their attention to evidence, and aesthetic in drawing conclusions from it. Many change their minds and paradigms only with great reluctance. The actual “culture of science,” in other words, is far from the white coat image that appears to the layperson. The accumulation of knowledge in actual science is not at all continuous—it moves by fits and starts. Real science is not about certainty but about uncertainty.

Indeed, contemporary philosophy of science is calling into question Hume’s “successionist” view of cause as involving regularities among repeated sequences of events—the notion that a single temporally antecedent phenomenon (which can be operationally defined as a unitary variable) causes a single consequent result, recurrently. In social analysis especially, it is apparent that, as Sayer observes, “the same causal power can produce different outcomes (for example, economic competition can prompt firms

Real science is not about certainty but about uncertainty.
to restructure and innovate or to close. Sometimes different causal mechanisms can produce the same result; for instance, you can lose your job for a variety of reasons" (Sayer, 2000, p. 15; see also Salmon, 1998). By this way of thinking, an adequate causal analysis must identify the specific mechanisms that generate specific outcomes within particular structural circumstances. It is just such specification that the Light and Pillemer (1982) example of schmoozing provides—a specification made possible by direct observation from within the local situation of complexity and contingency. The variety and changeability of the hierarchically embedded contexts of social life are such that simple, consistent associations between generic cause and generic effect of the sort tested in formal social experiments are not likely to occur. From this point of view the level of abstraction in the operational definition of aspects of social process as unitary variables—which characterizes large-scale social experiments—results in knowledge that at best can be characterized as rough approximation or guesswork. This is not to say that valid conclusions can never be drawn from experimental studies of “what works,” but rather that those who do such research need to cultivate skepticism and humility—a realistic assessment of the difficulties and limits that are inherent in their approach.

The Department of Education’s Strategic Plan holds up evidence-based medicine for emulation by educational research. And yet the moment one enters the arena of behavioral and social aspects of medical service delivery—the contingent circumstances of medical practice most similar to those of educational practice—one finds that the conclusions of experimental studies in medical research do not result directly in the outcomes of curing anticipated by researchers because of the enduring, embarrassing presence of locally constructed social facts (e.g., that many patients do not follow the prescribed medical regimen: They don’t take their pills or do their exercises). Qualitative researchers would say that these patients are making sense but that they don’t take their pills or do their exercises. Qualitative researchers would say that these patients are making sense but that they are inherently different from their approach.

Finally, although we have doubts about the appropriateness of treating the practice of medicine and education as analogous, let us stay with that analogy to make one more point. We are concerned that premature conclusions about “what works” in the short term, without careful consideration of side effects that may appear downstream, can provide false warrants for the educational equivalent of thalidomide. That was a medical treatment that was shown scientifically (i.e., by means of randomized trials) to have clear positive effects. A non-barbiturate hypnotic, thalidomide was originally prescribed after 1956 to prevent morning sickness in pregnant women and to help them sleep through the night. What a tragic irony: thalidomide prevented morning sickness very effectively but it was also effective in causing deformities in the fetus growing in the mother’s womb. The latter effects were only discovered after the babies were born, and it took years to trace the cause of the deformities back to the mothers’ use of thalidomide. Will our current desperate attempts to discover “what works” to raise standardized test scores in the short run have analogous effects on our children and teachers in school, effects that are only apparent after much damage has been done? How do we design experiments that can predict dangerous educational side effects?

To conclude, the unmixed optimism about science in educational research and the hopes for evidence-based social engineering as a means of educational reform appear to us to be unwise, even superstitious. In our experience, positive educational change is accomplished locally and it is more like walking through a swamp, testing the ground with each step, than it is like driving on a superhighway or even like building one. To get smarter about working our way in a swamp we need all kinds of research and deliberation, scientific and nonscientific. And we need practitioner research. The current federal discourse about improvement in educational research fails to emphasize practitioner research as a source of insight; yet if real progress is to occur in our swamp-situated efforts, it seems quite obvious to us that the knowledge of practitioners and of research specialists must grow together in new ways, a point made very recently in ER by Hiebert, Gallimore, and Stigler (2002).

Moreover, neither the Department of Education’s Strategic Plan nor the NRC report foregrounds the plain historical fact that educational research has been dramatically underfunded federally. If anywhere near the federal investment over the past century in military research and development, or that in medicine, had been provided for educational research, it is likely just on actuarial grounds that considerable insight would have been gained. For the federal government to criticize educational research for lack of progress while having long withheld adequate funding for it seems contradictory and even hypocritical.

We think it is deeply misleading to claim that expressways to educational progress can be engineered socially upon bedrock of smoothly accumulating and consistent evidence about effectiveness—and that this can be done readily on the cheap. The current federal leap of faith to science as a warrant for certainty in social policy reminds us of the jejune contempt held by the Renaissance for the Middle Ages, a “pessimism” that was repeated in the Enlightenment. For many at the beginning of the 20th century and at its ending as well, that hope in progress as the result of the continuous development of new knowledge came to appear as a delusion. Our concern can be summarized in a paraphrase of the old aphorism: Those who do not know their intellectual history are condemned to repeat it.

REFERENCES


**AUTHORS**

FREDERICK ERICKSON is the George F. Kneller Professor of Anthropology of Education in the Social Research Methodology Division, University of California, Los Angeles, Graduate School of Education and Information Studies, 1002 Moore Hall, Box 951521, Los Angeles, CA 90095-1521; ferickson@gseis.ucla.edu. He is also the Director of Research at the Corinne A. Seeds University Elementary School at UCLA. His research interests include issues of equity in urban education, the study of learning in social interaction, video-based analysis of oral discourse, and contemporary ethnographic research methods, including digital multimedia documentation of teaching practice.

KRIS GUTIERREZ is a full professor at the University of California, Los Angeles, Urban Schooling, Curriculum, Teaching, Leadership, and Policy Studies, 1026 Moore Hall, Los Angeles, CA 90095-1521; krisgu@ucla.edu. Her research interests focus on understanding the relationship between language, culture, and human development and include the study of the social and cognitive consequences of literacy, practices, and policies in urban schools. Her work also examines the effects of current educational reforms, including new language policies on poor and working-class students—particularly those who are English Language Learners.

Manuscript received June 10, 2002
Revisions received July 7, 2002
Accepted July 11, 2002

Curriculum Inquiry is dedicated to the study of educational research, development, evaluation, and theory. This leading international journal brings together influential academics and researchers from a variety of disciplines around the world to provide expert commentary and lively debate. Articles explore important ideas, issues, trends, and problems in education, and each issue also includes provocative and critically analytical editorials covering topics such as curriculum development, educational policy, and teacher education.

"One of the few education journals to open up contemporary theoretical perspectives on general education."

Maxine Greene, Columbia University, USA

A library subscription to the print volume entitles readers to:
- Free access to full text articles online
- Free copying for non-commercial course packs
- Free access to all available electronic back volumes

Select Blackwell Publishing's E-mail Updates

Free Curriculum Inquiry's Table of Contents e-mailed to you automatically!
To register for your FREE updates simply visit:
http://select.blackwellpublishers.co.uk

SELECT exactly what you want to receive
SELECT contents from the journals of your choice
SELECT news of books and journals by subject area
SELECT when your messages arrive—right down to the day of the week

Visit Curriculum Inquiry's website at: http://home.oise.utoronto.ca/~cli/