

Improving the Effectiveness of Education

The Need for Rigorous Research

Barbara Schneider

John A. Hannah Chair University Distinguished Professor
College of Education and Department of Sociology
Michigan State University



In the press release announcing the award of the Regional Educational Laboratories contracts, Russ Whitehurst is quoted as saying “We look forward to a very productive relationship as we work to meet the research needs of those engaged in improving the effectiveness of education.” To do this the labs need to ‘... work to ensure that those involved in educational improvement at the local, state, and regional levels have access to the best available information from research and practice.’ (See the March 28, 2006 U.S. Department of Education press release “Regional Educational Lab Contracts Awarded: Labs to provide states and educators with best research and practices to improve education,” available from <http://www.ed.gov/news/pressreleases/2006/03/03282006.html>; link verified May 5th 2006.)

NOTE: Included in the notes sections on the following pages is material extracted from the forthcoming (2006) Schneider et al *Estimating Causal Effects Using Experimental and Observational Designs: A Think Tank White Paper*. Please contact the author directly for permission to quote, cite, or reproduce this material.

How do we obtain the “Best available information”?



- Need evidence that warrants action.

For example:

- Which curricula, pedagogy, interventions should we adopt?
- How should resources be expended?

What kind of evidence warrants action?



- Evidence based on scientific research principles
 - See, e.g., *Scientific Research in Education* (2002):
 - Pose significant questions that can be investigated empirically
 - Link research to relevant theory
 - Use methods that permit direct investigation of the question
 - Provide a coherent and explicit chain of reasoning
 - Replicate and generalize across studies
 - Disclose research to encourage professional scrutiny and critique

But fundamentally [ON NEXT SLIDE] research questions determine the most appropriate designs and methods



Research questions determine the most appropriate designs and methods

What are the incentives for evidence-based education ?



- Recent legislation
 - *No Child Left Behind Act of 2001* (PL 107-110)
 - *Education Sciences Reform Act of 2002* (PL 107-279)
- Funding priorities
 - RFAs & program announcements (IES, NSF, NICHD)

“... the No Child Left Behind Act of 2001 (NCLB) provided a specific definition of scientifically-based research and set aside funding for educational research studies consistent with this definition. ... NCLB places special emphasis on determining which educational programs and practices are effective through rigorous scientific study. This concept was reinforced through The Education Sciences Reform Act of 2002, which replaced the Office of Educational Research and Improvement (OERI) with the newly created Institute of Education Sciences (IES). ... This legislation states that by conducting scientifically-based research studies that apply rigorous, systematic, and objective methodology to obtain reliable and valid knowledge, it is possible to identify educational practices and activities that result in improved student learning. These goals were reiterated in a statement by the Secretary of Education in the January 25, 2005 *Federal Register*, where it is noted that ‘random assignment and quasi-experimental designs [are considered] to be the most rigorous methods to address the question of project effectiveness’ (p. 3586),” (Schneider et al., forthcoming, 2006). (See also Department of Education RIN 1890-ZA00, “Scientifically Based Evaluation Methods; Notice,” *Federal Register* 70(15) 3586-3589.)

“... funded research programs in the Institute of Educational Sciences (IES), the National Science Foundation (NSF), and National Institute for Child Health and Human Development (NICHD) have [also] increased calls for intervention studies that provide clear evidence of student learning,” (Schneider et al., forthcoming, 2006).

Responses by the educational research community to evidence-based education



- Provide guidelines in journal reporting
- Causal report on estimating causal effects and experimental observational designs
- Redesign of doctoral training programs
- Recent initiatives of the National Research Council (NRC)

Ways of learning what works



- Use of research designs that are appropriate for making causal inferences (*e.g.*, experimental or quasi-experimental designs)
- Value of descriptive information for policy purposes

“... questions of causality [are] at the forefront of [current] educational debates and discussions in part because of dissatisfaction with the quality of educational research and recent federal initiatives designed to promote the accumulation of scientific evidence in education that rely on randomized controlled trials (RCTs). A common concern expressed by those deeply engaged with the educational enterprise, as well as those outside of education, revolves around the design of and methods used in educational research, which many claim have resulted in fragmented and often unreliable findings. ...Pointing to lack of replication, inappropriate designs for assessing causal effects, and crude analytic procedures, some researchers have argued that it is difficult to accumulate a knowledge base that has value for practice or future study,” (Schneider et al., forthcoming, 2006).

Gold standard for estimating effects



- Randomized controlled trials (RCTs)
 - **BENEFITS** of RCTs
 - Eliminate selection bias – no systematic differences in pre-treatment characteristics between treatment and control groups

“... when financially, logistically, and ethically feasible, the randomized field trial is the best design for making causal inferences about the effectiveness of educational programs and practices,” (Schneider et al., forthcoming, 2006).

“The random assignment of participants to treatment conditions maximizes the probability that treatment group assignment is independent of the pretreatment characteristics of group members; thus differences between groups can be attributed to treatment effects rather than these pretreatment characteristics.” Of course “[r]andomized experiments ... indicate only whether there are treatment effects and the magnitude of those effects; they do not identify the mechanisms (i.e., the specific aspects of the treatments in question or the settings in which they are implemented) that may be contributing to such effects.” Still, “correctly implemented,” RCTs are “the most powerful designs for detecting treatment effects,” (Schneider et al., forthcoming, 2006).

Gold standard for estimating effects *(continued)*



- **HOWEVER**, RCTs are not always feasible logistically or ethically
- **SO**, we need alternatives, *e.g.*,
 - Quasi-experimental designs
 - Use of large-scale observational studies to approximate RCTs

Secondary analyses of large-scale datasets



- Benefits of using large-scale, representative data sets:
 - Generalizable to specific populations of students
 - Rich source of descriptive information
 - Useful in developing plausible causal hypotheses
 - Provide preliminary results to inform RCTs

How do we estimate causal effects using observational data?



- *What are the challenges?*
 - Selection bias
 - Omitted variables
- *How can they be addressed?*
 - Fixed effects models
 - Instrumental variables
 - Propensity score matching
 - Regression discontinuity

Examples of secondary analysis for policy purposes



- Hong & Raudenbush:
 - Retention in kindergarten – no evidence improves average achievement in math or reading
- Nye, Konstantopoulos, & Hedges:
 - Teacher effects – quality matters
- Downey, von Hippel, & Broh:
 - Value of summer learning
- Cameron & Heckman:
 - Value of the high school diploma in contrast to GED

Hong, Guanglei, & Raudenbush, Stephen W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27(3): 205-224. Available online from http://www.aera.net/uploadedFiles/News_Media/News_Releases/2005/Kindergarten%20Retention-Hong%20&%20Raudenbush%20PDF.pdf (verified May 4, 2006).

NOTE: Also available online is the related American Educational Research Association (AERA) News Release "Scholars Find Kindergarten Retention Fails to Help – and May Hinder – Academic Achievement in Reading and Mathematics" at <http://www.aera.net/newsmedia/?id=889> (verified May 4, 2006).

Nye, Barbara, Konstantopoulos, Spyros, & Hedges, Larry V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257.

NOTE: An abstract of this article is available online at <http://www.aera.net/publications/?id=326> (verified May 4, 2006).

Downey, Douglas B., von Hippel, Paul T., & Broh, Beckett A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review*, 69: 613-635.

Cameron, Stephen V., & Heckman, James J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics* 11(1) Part 1: Essays in honor of Jacob Mincer: 1-47.

Bringing awareness to a local problem



A principal working with Census data was able to show diversity in the school population and the surrounding neighborhood suggesting a disconnect between the cultural backgrounds of the students and the perceptions of their teachers.

What we would like to do: Leverage information



- A wealth of data exists at the school, district, state, and federal levels that should in principle provide an empirical basis for developing educational policies, practices, and intervention effectiveness studies
- These data sets are very costly endeavors
- We need collaboratively to develop more systematic ways to help those in the MREL region use these data to guide policymaking and practice

Leveraging information in the MREL region: An approach



First Step:

***Identify state, district, and local
Information needs***

Our approach for leveraging information in the MREL region



- Use multiple levels of data to address key issues in the region by:
 - Developing an inventory and index of available data
 - Linking and preparing customized datasets
 - Recruiting and training staff with skills required to conduct innovative and appropriate analyses of large-scale observational datasets