

Narrow Questions, Narrow Answers: The Limited Value of Randomized Controlled Trials for Education Research

ANNETTE LAREAU*

With their emphasis on “scientific rigor” and references to a “gold standard,” recent federal guidelines leave little room for debate about the importance of randomized controlled trials for education research.¹ In this chapter I offer a critical assessment of this position. I suggest that randomized controlled trials, although useful in selective research situations, are not a helpful research method for the field as a whole. The most important questions in education simply cannot be answered with this approach. I argue that the medical research model has been misunderstood by advocates of randomized controlled trials in education and misapplied to the field of education research.

The limited value of randomized controlled trials in education, and in various areas of medicine, has many sources. I focus on three. One problem is that the research questions that can be asked using this method are relatively narrow and thus cannot accommodate the complexity of the educational process. The method seeks to control all features except for one important variable or “treatment” effect. Yet, as I argue below, educational institutions are multifaceted organizations; there are constant changes in many of the “independent” variables that have an impact on the “dependent” or outcome variables. A further aspect of the complexity problem is that it involves questions of meaning. Randomized controlled trials cannot address this type of “variable.” Understanding organizational processes and the meanings participants assign to events is crucial to the development of a rich picture of educational institutions; the effects of educational treatments or variables depend in critical ways on what they mean to students, parents, and teachers, among others. Important research questions—such as the mechanisms through which the position of parents transmits advantages, the processes through which some principals are

more effective than others, or the micro-interactive strategies that build trust among principals and teachers—thus fall outside the methodological realm of randomized controlled trials.

Second, randomized controlled trials pose formidable problems of execution. It is common for researchers to have difficulty selecting a control group that truly acts as a control. In education, even studies blessed with talented staff and generous resources have encountered difficult problems. Researchers cannot sustain core elements of the method, including the expectation that treatment groups are delivered the treatment and control groups do not receive it. Also, randomized controlled trials are based on the assumption that the control group is deprived of the potentially valuable treatment. In a population of schoolchildren, this traditional assumption raises complex ethical concerns. Researchers promoting randomized controlled trials are guilty of a naive hopefulness that these problems in execution will not overwhelm the research. Yet previous studies (some of which I discuss in this chapter) have not been promising in terms of the ability of researchers to execute core elements of the design.

Third, in their enthusiastic embrace of randomized controlled trials as an education research method, policy analysts have not paid sufficient attention to the crucial issue of implementing research results. Decades of research in education clearly demonstrate that it is not possible for policymakers to mandate the successful adoption of an educational policy. Rather, successful reform requires that educators “buy into” the new policy. In medicine, too, “buy-in” significantly affects the likelihood of achieving change. Even when medical research demonstrates that many lives and millions of dollars can be saved by a relatively simple innovation, the necessary change may be resisted if it runs counter to the logic and culture of the organization. Moreover, as I discuss below, in both education and medicine, promising reforms may be difficult or impossible to implement or “scale up” beyond the initial research setting.

In sum, there are formidable limits to the randomized control approach—limits which the proponents of this methodology have not sufficiently recognized. To achieve “rigor” in education research, we need to adopt a more realistic grasp of the problems that plague attempts to implement research results as well as a greater openness to other approaches that offer crucial insights into social processes in organizations. For these purposes, qualitative methods, including participant observation and in-depth interviews, are likely to be the most promising.

What Kinds of Questions Can Be Answered?

Randomized controlled trials involve the random assignment of participants to different experiences or “treatments,” with all other components of the experience “held constant” (Mosteller and Boruch 2002). A close cousin to a randomized control trial is a study that compares individuals who received a

“treatment” with individuals who sou the waiting list. As the materials in demonstrate, randomized controlled recent years. The authorizing legislati the importance of using “scientifically A). Cook and Payne (2002:177) provi

The superiority of random as about the consequences of plat acknowledged in philosophy, m statistics, micro-economics, psy research, early childhood educat political science and sociology c surveys. . . . This article does not a come only from experiments. It d provide a better warrant for such . So, if experiments can be conduct use them requires a very strong ja

It is relatively rare that social scient controlled trial in a high-quality, thoug they do, the results can be crisp, clear. trial is appropriately designed and co fically to evaluate the effect of an int be stunningly appealing. One exampl Sister program in Philadelphia, cond Private Ventures (Tierney, Grossman, those on the waiting list for a Big Broo program. The researchers found that) compared to those on the waiting list, been effectively handled in this design, has a positive impact. Similarly, in a se Opportunity and New Hope reforms a low-income families to rent an apartm ethnographic research approach (Dun Liebeman, and Katz 2007). This rando the housing service that placed fam neighborhoods improved mothers' m depression. The housing service also ha not on other measures of physical hea raise adults' income, however, and the were mixed: generally quite positive for et al. 2007). Importantly for my purpos itself could not shed light on why effect

to-interactive strategies that build trust fall outside the methodological realm of

als pose formidable problems of execution. It is difficult to select a control group, even studies blessed with talented staff encounter difficult problems. Researchers' method, including the expectation that treatment and control groups do not receive the valuable treatment. In a population intervention raises complex ethical concerns. Randomized controlled trials are guilty of a naive expectation will not overwhelm the research. (I discuss in this chapter) have not been able to execute core elements of the

of randomized controlled trials as analysts have not paid sufficient attention to research results. Decades of research in it is not possible for policymakers to implement educational policy. Rather, successful implementation of the new policy. In medicine, too, "buying" of achieving change. Even when medical change and millions of dollars can be saved by necessary change may be resisted if it runs counter to the organization. Moreover, as I discuss elsewhere, promising reforms may be difficult or impossible beyond the initial research setting.

(to the randomized control approach—methodology have not sufficiently recognized. In research, we need to adopt a more realistic approach to implement research results as approaches that offer crucial insights into these purposes, qualitative methods, and in-depth interviews, are likely to be the

answered?

the random assignment of participants to the treatment group with all other components of the program (Cook and Boruch 2002). A close cousin to a randomized trial is a study that compares individuals who received a

"treatment" with individuals who sought the same treatment but remained on the waiting list. As the materials in the appendices to this volume clearly demonstrate, randomized controlled trials have skyrocketed in popularity in recent years. The authorizing legislation for No Child Left Behind emphasizes the importance of using "scientifically based research" methods (see Appendix A). Cook and Payne (2002:177) provide a ringing endorsement:

The superiority of random assignment for drawing inferences about the consequences of planned change attempts is routinely acknowledged in philosophy, medicine, public health, agriculture, statistics, micro-economics, psychology, criminology, prevention research, early childhood education, marketing and those parts of political science and sociology concerned with improving opinion surveys. . . . This article does not argue that correct causal conclusions come only from experiments. It does argue, though, that experiments provide a better warrant for such conclusions than any other method. So, if experiments can be conducted in schools, they should be. Not to use them requires a very strong justification.

It is relatively rare that social scientists successfully carry out a randomized controlled trial in a high-quality, thoughtful, and well-executed study. But when they do, the results can be crisp, clear, and causally irrefutable. Indeed, when a trial is appropriately designed and conducted, and when its purpose is specifically to evaluate the effect of an intervention or treatment, the results can be stunningly appealing. One example is an evaluation of the Big Brother/Big Sister program in Philadelphia, conducted by the non-profit group Public/Private Ventures (Tierney, Grossman, and Resch 2000). The study compared those on the waiting list for a Big Brother/Big Sister with those involved in the program. The researchers found that program participants had less drug use compared to those on the waiting list. Since the problem of self-selection has been effectively handled in this design, the results are convincing: The program has a positive impact. Similarly, in a series of thoughtful studies the Moving to Opportunity and New Hope reforms used a random assignment of services for low-income families to rent an apartment with a combined quantitative and ethnographic research approach (Duncan, Huston, and Weisner 2007; Kling, Liebman, and Katz 2007). This random-assignment study clearly showed that the housing service that placed families in less economically depressed neighborhoods improved mothers' mental health by reducing their rates of depression. The housing service also had an impact on participants' obesity, but not on other measures of physical health. The neighborhood context did not raise adults' income, however, and the results for children's school performance were mixed: generally quite positive for girls but quite negative for boys (Kling et al. 2007). Importantly for my purposes here, the randomized controlled trial itself could not shed light on *why* effects differed for girls and boys. For insights

into why gender might have had an impact, the study team turned, appropriately, to ethnographic research (Clampet-Lundquist et al. 2006).

Even when conducted in tandem with qualitative investigations that are better suited to answering why and how questions, random-assignment studies have formidable built-in constraints. They must isolate one dependent variable and control all other factors except the "treatment." Examination of interaction effects (i.e., where the results are contingent upon one set of circumstances but not another) is possible, but cumbersome, as Cook and Payne (2002:152) acknowledge:

At their most comprehensive, [random-assignment] experiments can responsibly test only a modest number of the possible interactions between treatments. So, experiments are best when a causal question involves few variables, is sharply focused, and is easily justified.

The models also presume that during the study period, conditions are static, not dynamic. Educational institutions do not conform to these basic assumptions. Schools are complex organizations with many different levels, including the classroom (and student groupings within the classroom), the school, and the district. Many of these levels are in flux, particularly in a climate of educational reform (Hubbard, Stein, and Mehan 2006; Sunderman, Kim, and Orfield 2005). Dynamics in classroom learning are shaped by many factors; it is difficult to delineate sharply one causal question.

Randomized controlled trials focus on factors that can be manipulated via policy intervention. But crucial aspects of social life that affect education cannot be manipulated this way. For example, there is an extensive body of research on the friendship networks of youth and the influence of peers on educational aspirations, homework patterns, and other aspects of schooling (Hallinan 2006). Students cannot be assigned to different friendship networks by a research study, let alone be assigned randomly to peer groups by researchers. Similarly, social-class differences in child rearing, in parents' knowledge about educational institutions, and in the flexibility parents have for attending school events during work hours are not readily subjected to manipulation (Entwisle, Alexander, and Olson 1997; Lareau 2000, 2003). Arguably, many classic studies in education would never have been undertaken in the current research climate, since the topics these studies have addressed—peer networks, physical attractiveness, immigrant status, and students' aspirations—would not be amenable to a randomized controlled trial approach. Further, decades of research on educational achievement have shown that it is the elements least open to external manipulation (especially family background and peer networks) that are the most important influences on student learning (Hallinan 2006). In randomized controlled trials the scope of questions that can be studied is narrowed considerably, leaving unstudied precisely the factors that prior research has identified as key influences on educational outcomes.³

The "narrow questions/narrow answers" approach to research in health and social conditions have yielded many important findings. But randomized controlled trials, arguably the most important methodological tool in health research, face a critical problem: that many patients do not comply with the treatments specified. (Haynes, McDonald, and Garg 2002: 152 and Education 2007). Research suggests that many patients are non-compliant (or non-adherent) with treatments at the time(s) specified.⁴ Non-compliance is a problem for psychological disorders (e.g., anorexia nervosa, patients with manic depression) (O'Connell, Patient Communication Problems and Their Medical Problem also are ubiquitous) (Information and Education 2007). For example, employers, the difficulty of purchasing affordable insurance, and the burden on residents shouldered compared to the burden among today's most important policy issues through research based on the method of Rubin (2007).

Nor is random assignment useful in cases where the event is a result of malfunction in several institutions. For example, the case with natural disasters. For example, the multiple related deaths in Chicago during one of the worst heat waves found that multiple social institutions affected the death of the elderly: the para-military services, the social isolation of the elderly in the apartments and buildings housing the elderly during the health-care crisis by political organizations, and patterns resurfaced in the Hurricane Katrina disaster. The nature as well as the multiple institutional factors do not permit study via the method of random assignment (1978; Klinenberg 2002). Similarly, social events affected by multiple institutions, of which the method of random assignment is not applicable.

Moreover, the randomized controlled trial approach is not applicable to events in the lives of individuals. Take, for example, the research that has been done on parent involvement in education. It is little doubt that teachers value parent involvement. They call for "partnerships" in education. Parents want to take an active role in their children's education. There have been interventions to increase parent involvement was promoted in

fact, the study team turned, appropriately, to qualitative investigations that are

with qualitative investigations that are narrow questions, random-assignment studies. They must isolate one dependent variable and one "treatment." Examination of interaction is dependent upon one set of circumstances but not, as Cook and Payne (2002:152)

random-assignment] experiments can handle a number of the possible interactions. Experiments are best when a causal question is focused, and is easily justified.

During the study period, conditions are static, and do not conform to these basic assumptions with many different levels, including within the classroom), the school, and in flux, particularly in a climate of educational change (Mehan 2006; Sunderman, Kim, and others). Learning are shaped by many factors; it is a complex question.

Factors on factors that can be manipulated via the social life that affect education cannot be studied, there is an extensive body of research on the influence of peers on educational achievement and other aspects of schooling (Hallinan 2006). Different friendship networks by a randomly to peer groups by researchers. Child rearing, in parents' knowledge about child development, ability parents have for attending school, and ability subjected to manipulation (Entwisle, 2003). Arguably, many classic studies undertaken in the current research climate, including—peer networks, physical attractiveness—peer networks, physical attractiveness—peer networks, physical attractiveness—peer networks—would not be amenable to random assignment. Further, decades of research on parent involvement that it is the elements least open to random assignment (background and peer networks) that affect student learning (Hallinan 2006). In the face of questions that can be studied precisely the factors that prior research has identified on educational outcomes.²

The "narrow questions/narrow answers" problem is also found in medicine. For example, drug trials with a double-blind design, a placebo, and controlled conditions have yielded many important insights about which drugs are most effective.³ But randomized controlled trials have not provided insight into arguably the most important questions in medicine today. For example, a critical problem is that many patients are non-compliant in the taking of drugs (Haynes, McDonald, and Garg 2002; National Council on Patient Information and Education 2007). Research suggests that approximately one-half of the time patients are non-compliant (or non-adherent) in taking pills or other medical treatments at the time(s) specified.⁴ Non-compliance rates are particularly high for psychological disorders (e.g., an 80% rate of non-compliance among patients with manic depression) (Osterberg and Blaschke 2005). Doctor-patient communication problems and patients' failure to grasp the nature of their medical problem also are ubiquitous (National Council on Patient Information and Education 2007). The high cost of medical insurance for employers, the difficulty individuals with chronic health problems have purchasing affordable insurance, and the relatively large medical costs U.S. residents shoulder compared to their Western European counterparts are among today's most important policy issues, yet they cannot be addressed through research based on the method of random assignment (Furman and Rubin 2007).

Nor is random assignment useful in understanding health crises that occur as a result of malfunction in several institutions at the same time, as is often the case with natural disasters. For example, in his analysis of the hundreds of heat-related deaths in Chicago during one week in 1995, Eric Klinenberg (2002) found that multiple social institutional factors contributed to the high rates of death of the elderly: the para-military organization of the emergency health services, the social isolation of the elderly, lack of cooling facilities in the apartments and buildings housing the elderly, and the slow identification of the health-care crisis by political organizations and the media. Many of these patterns resurfaced in the Hurricane Katrina disaster. These disasters' rare nature as well as the multiple institutions involved in shaping the (ineffectual) response do not permit study via the method of random assignment (Ericksen 1978; Klinenberg 2002). Similarly, schooling is a complicated process that is affected by multiple institutions, of which only one is schools themselves.

Moreover, the randomized control trial gives short shrift to the *meaning* of events in the lives of individuals. Take, for example, the vast amount of research that has been done on parent involvement in schooling—research that identifies parent involvement as an important influence on children's achievement. There is little doubt that teachers value parent involvement highly (Epstein 2002). They call for "partnerships" in education with parents. Most surveys also reveal that parents want to take an active, helpful role in their children's schooling.⁵ There have been interventions to increase parent involvement in schooling; parent involvement was promoted in the What Works Clearinghouse (U.S.

Department of Education 2006). But these studies do not sufficiently recognize that social class appears to influence what parents *mean* by the term "parent involvement." For example, in selecting schools, middle-class parents appear to collect very detailed information about educators, and these parents oversee their children's educational lives closely (Diamond and Gomez 2004; Lareau 2000). Working-class parents, on the other hand, also see themselves as heavily involved in schooling, but they often interpret involvement to mean preparing children to go to school and deferring to the educational professional expertise of educators (Lareau 2000, 2003).

It is important to remember that the study of what people do is only one piece of the puzzle. Complex social processes undergird educational reform. Small, intensive, non-random case studies are crucial for identifying these processes and shedding light on the conditions associated with successful reform initiatives.

Barriers to Randomized Controlled Trials

In practice, the emphasis on "rigor" incorporated into No Child Left Behind has come to be equated with a commitment to and reliance on randomized controlled trials (see the appendices to this volume). Indeed, Institute of Education Sciences Director Grover "Russ" Whitehurst has termed randomized controlled trials the "gold standard." The approach has many advocates. Yet despite the desire of researchers to implement a medical model, there are problems. One is difficulty in selecting and sustaining control groups.

When Is a Control a Control?

Sustaining the quality and quantity of the control group is challenging, as the authors of a review of diabetes research note (Montori et al. 2006):

In conducting systematic reviews of RCTs [randomized controlled trials] in diabetes we have noticed that researchers seem to pay little attention . . . to methodological safeguards that limit the introduction of bias into RCTs. As a result, these potentially biased RCTs could mislead clinicians. . . . When reports leave out critical information about methodological safeguards against bias, readers cannot ascertain if these safeguards were present.

(p. 1833)

The authors reviewed RCTs published in general medical journals (e.g., the *New England Journal of Medicine*), in those that specialize in diabetes research (e.g., *Diabetes*), and in those that address metabolics and nutrition (e.g., *American Journal of Clinical Nutrition*). They described their evaluation criteria this way:

We considered trials to be of low methodological quality when they had three or more of these criteria: inadequate (or not reported)

allocation concealment, inadequate blinding of patients and caregivers, failure to follow the randomization principle, or a reporting >10% (or not calculated) of loss to follow-up.

The authors judged 53% of the 11 studies to be of low quality. That figure is itself a cause for concern, as it is likely an underestimate. They focused on a review of studies across a broad range of settings, which may show a higher rate of low-quality studies.

Similar quality-related problems arise from the most basic assumption of randomized controlled trials: that the experimental group and the experimental/treatment group will not receive the treatment "dosage" while the control group will not receive the treatment "dosage" will be a problem for those who have attempted to use randomized controlled trials.

When educational reforms are implemented elsewhere. For example, in promoting education of at-risk youth, researchers sought to improve education through parent involvement programs, and this highly regarded reform into a highly regarded reform into an unusually elaborate array of randomized controlled trials. Implementing a study based on randomized controlled trials is particularly difficult in a tumultuous environment. It was experiencing district politics, changes in administration, and changes in student population.

Thomas Cook was among those who reported the results of this research. He explicitly acknowledged some of the difficulties of the randomized control approach to evaluate a development program:

Unfortunately 5 of the 24 schools were dropped from the study at various times for various reasons. . . . Some schools were not appointed who did not want to be part of the study. Some did not want the quantitative approach. Some schools stayed in the program but were not included (continued). Of the Phase II schools

ese studies do not sufficiently recognize what parents *mean* by the term "parent" (schools, middle-class parents appear to be educators, and these parents oversee them). On the other hand, also see themselves as heavily involved in preparing the child for the educational professional expertise

he study of what people do is only one of the processes undergirding educational reform. These studies are crucial for identifying these conditions associated with successful

Trials

incorporated into No Child Left Behind (commitment to and reliance on randomized control trials to this volume). Indeed, Institute of Education's Whitehurst has termed randomized control trials "The approach has many advocates. To implement a medical model, there are control and sustaining control groups.

the control group is challenging, as the authors note (Montori et al. 2006):

of RCTs [randomized controlled trials] that researchers seem to pay little heed to the safeguards that limit the introduction of potentially biased RCTs could result in trials that leave out critical information and, against bias, readers cannot ascertain

(p. 1833)

in general medical journals (e.g., the *New England Journal of Medicine*) that specialize in diabetes research (e.g., *Diabetes Care*), and in pediatrics and nutrition (e.g., *American Journal of Clinical Nutrition*) attributed their evaluation criteria this way:

methodological quality when they were inadequate (or not reported)

allocation concealment, inadequate (or not reported) blinding of patients and caregivers, failure to adhere to the intention-to-treat principle, or a reporting >10% (or did not report information to calculate) of loss to follow-up.

(p. 1834)

The authors judged 53% of the 199 trials they evaluated to be of "low quality." That figure is itself a cause for concern, but the authors suggest that it may be an underestimate. They focused only on work published in "top journals"; a review of studies across a broader sample of scientific journals would likely show a higher rate of low-quality studies (Montori et al. 2006).

Similar quality-related problems have surfaced in education studies. Indeed, among the most basic assumptions of any randomized trial are that the control group and the experimental/treatment group are highly similar before the start of the study, that the experimental group will receive a full course of treatment, while the control group will not receive *any* exposure to the treatment, and that the treatment "dosage" will be delivered consistently over time. In school settings these assumptions are difficult or impossible to meet, as researchers who have attempted to use randomized controlled trials have discovered.

When educational reforms prove successful in specific settings, they often are implemented elsewhere. For example, James Comer's complex reform for promoting education of at-risk youth was adopted in Chicago. The program sought to improve education through community involvement, intensive parent involvement programs, and community programs. The introduction of this highly regarded reform into the Chicago public schools was subject to an unusually elaborate array of randomized controls and program evaluation. Implementing a study based on randomized control methodology is exceptionally difficult in a tumultuous organizational environment, however. At the start of the reform effort, the Chicago public school system was just such an environment. It was experiencing high rates of teacher turnover, changing district politics, changes in administrator leadership, and turnover in the student population.

Thomas Cook was among those studying the reform. In published papers reporting the results of this research, he and his colleagues (to their credit) explicitly acknowledge some of the barriers they faced trying to use a randomized control approach to evaluating the impact of the Comer school development program:

Unfortunately 5 of the 24 schools dropped out of the study at different times for various reasons. . . . In two cases, a new principal was appointed who did not want the program; in the third, the principal did not want the quantitative research component (although his school stayed in the program and the ethnographic component continued). Of the Phase II schools, only one dropped out, [but it was]

almost immediately after learning it had been assigned to control status. . . .

(Cook, Murphy, and Hunt 2000: 544)

Such selective attrition vitiates the randomized experiment because more Comer than control schools dropped out and because the treatment is clearly confounded with principal turnover.

(pp. 544–545)

That “a new principal . . . did not want the program” and that one-fifth of the schools dropped out of the study for periods of time are critical problems. Other difficulties Cook and his co-authors do not mention include researchers’ lack of control over subject turnover (in addition to student and teacher turnover, top leadership at participating schools can change); and researchers’ limited power either to shore up principals’ flagging interest in participating in a study or to overcome their initial resistance to such participation. As a result, randomized controlled trials that focus on longitudinal change in schools face daunting, and arguably insurmountable, challenges.

Of even greater concern is the fact that the double-blind character of the medical model, made possible by a clinical environment in which a treatment pill and a “sugar pill” look identical, has no counterpart in education. In the social and political world of schools, there is no way to create a double-blind reform. School reforms are discussed by teachers, principals, district leaders, and, in some schools, by parents and students as well. These parties do not live in a vacuum, and educators in the “control” and the “experimental” schools have social and professional relations. As a result, studies are easily contaminated: In a real-world educational setting, it is not possible for researchers to prevent the control group from being exposed to the treatment. As Cook et al. (1999) acknowledge in the Chicago Comer reform study:

Three comparison schools borrowed some program elements: In one case, a husband and wife worked in different schools, one with and one without the program, in another case the daughter of a senior New Haven program central staff member worked as an administrator at a control school; and, in the last case, a comparison group principal liked the program, studied it for himself, and discussed it with principals from program schools. And further diffusion between program and control schools probably occurred during district-wide in-service training sessions conducted by the county program coordinator.

(p. 584)

The fact that a principal in a “control” school liked the program and the program was diffused in district-wide training sessions is a sign of the social character of education. Still, Cook and his co-authors discount the impact of this source of contamination on their results:

But we judge the degree of diffusion to be relatively minor. The three schools were the major program elements in the school, namely the in-school training sessions, and training sessions were reclassified as control schools, so reclassifying these three control schools would have little effect on the estimate of student effects.

But it is clear that the most basic element of the experimental design—random assignment—was not adequately met. Nor could “control” schools be consistently delivered.

Midway through the study the principal of the control school and his appointees . . . reduced the intensity of the program. The new emphasis made test scores a more important measure of school effectiveness (rather than student learning and management). Because of the changes, the control schools were put on probation.

Hence, the vigor with which basic research is conducted across time. This pattern has been repeated in many studies.

In sum, medical research shows that subjects in “treatment” and “control” groups are not equivalent. Large numbers of subjects are lost to follow-up. The subjects are vastly more complex and “noisy” than in a laboratory. The evaluation had an ethnographic component. “[Before the end of the study] the control schools were not any of the programs as faithfully followed, although some were very close” (Cook et al. 1999). If being faithfully followed, then it is not surprising that the effectiveness was lower.

Translating Legislation into Practice: Implementation

Finally, it is worrisome that the process of educational reform have paid so little attention to the implementation of intervention. There is ample evidence of policy and its implementation. (Mehan 2006; McLaughlin 1987). As pointed out:

ing it had been assigned to control

(Cook, Murphy, and Hunt 2000: 544)

the randomized experiment because
schools dropped out and because the
high principal turnover.

(pp. 544–545)

of the program” and that one-fifth of the
schools of time are critical problems. Other
do not mention include researchers’ lack
of attention to student and teacher turnover,
which can change); and researchers’ limited
ongoing interest in participating in a study
due to such participation. As a result,
even on longitudinal change in schools face
these challenges.

that the double-blind character of the
medical environment in which a treatment
has no counterpart in education. In the
there is no way to create a double-blind
study by teachers, principals, district leaders,
students as well. These parties do not live
in “control” and the “experimental” schools
separately. As a result, studies are easily con-
founded. In setting, it is not possible for researchers
to be exposed to the treatment. As Cook et
al. (2000) reform study:

covered some program elements: In one
of the different schools, one with and one
without the program, the daughter of a senior New
Haven teacher worked as an administrator at a
school. A comparison group principal liked
the program, and discussed it with principals
in the district. This diffusion between program and
control schools occurred during district-wide in-service
sessions and during county program coordinator.

(p. 584)

of school liked the program and the
in-service training sessions is a sign of the social
diffusion. His co-authors discount the impact of
these results:

But we judge the degree of diffusion from these sources to have been
relatively minor. The three schools in question did not have access to
the major program elements that build and sustain the program in a
school, namely the in-school facilitator, local program in-service
training sessions, and training sessions in New Haven. Moreover,
reclassifying these three control schools made no difference in analysis
of student effects.

(p. 584)

But it is clear that the most basic element of the design, notably that there would
be a distinctly different experience for “control” and “experimental” groups,
was not adequately met. Nor could researchers be confident that the “treat-
ment” would be consistently delivered to the treatment group:

Midway through the study the political context changed. The mayor
and his appointees . . . reduced the power of local school councils. . . .
The new emphasis made test scores the dominant criterion for judging
school effectiveness (rather than success in creating decentralized
management). Because of their low test scores, one sixth of all city
schools were put on probation.

(Cook et al. 2000: 541)

Hence, the vigor with which *basic elements* of the reform were enacted differed
across time. This pattern has been found in many other studies.

In sum, medical research shows that it is difficult to sustain trials where
subjects in “treatment” and “control” groups have radically different experi-
ences. Large numbers of subjects are frequently non-compliant. School settings
are vastly more complex and “noisy” than medical studies. Indeed, the Cook
evaluation had an ethnographic component which offered a powerful critique:
“[Before the end of the study] the ethnographers were not willing to classify
any of the programs as faithfully following all of the program guidelines
although some were very close” (Cook et al. 2000: 564). If the programs are not
being faithfully followed, then it is difficult to accurately evaluate their
effectiveness.

Translating Legislation into Practice: The Problem of Implementation

Finally, it is worrisome that the proponents of randomized controlled studies
of educational reform have paid so little attention to the nature of the policy
intervention. There is ample evidence that the road between the passage
of policy and its implementation is long and rocky (Hubbard, Stein, and
Mehan 2006; McLaughlin 1987). As Milbrey McLaughlin (1987: 172) has
pointed out:

[T]he overarching, obvious conclusion running through empirical research on policy implementation is that it is incredibly hard to make something happen. . . . It's hard to make something happen primarily because policymakers can't mandate what matters. We have learned that policy success depends critically on two broad factors: local capacity and will. . . . But will, or the attitudes, motivation, and beliefs that underlie an implementer's response to a policy's goals or strategies, is less amenable to policy intervention.

"Will" cannot be mandated but it plays a crucial role in implementation. Yet advocates of randomized controlled trials have failed to come to terms with the fact that the factors that can easily be manipulated experimentally, and thus whose effects can be studied in a randomized controlled trial, may be precisely those elements that matter the least in the grand scheme of schooling and learning.

Literature on implementation of public policy also suggests that local context significantly shapes implementation. For example, in the randomized trial evaluation of the Comer School Development Program researchers were unable to achieve a standardized implementation. Instead, the reform varied to "maximize the fit to local circumstances" in Chicago. Cook explains:

Elements of the implementation of [Comer's] SDP [School Development Program] vary by both district and school in order to maximize the fit to local circumstances. The Chicago program is unique in many ways. It was introduced in the middle of a citywide educational reform initiated by local politicians, businesses and philanthropies. . . . At first, the reform emphasized local school councils; . . . councils were encouraged to set their own school goals and establish their own instructional designs . . . [although this pattern] . . . did not arise in practice.

(Cook et al. 2000: 541)

If policy implementation is not standardized, then the validity of an assessment of its impact using a randomized controlled trial is compromised.

Several studies suggest that principals play a critical role in policy implementation. One study of "teacher accountability" in California reported:

Our research revealed that principals often played a pivotal role in how teachers experienced accountability policies. Principals either acted as a "buffer," shielding teachers from test-score pressures, or as an added source of pressure for teachers, emphasizing the need to raise test scores and the school's ranking.

(PACE 2006: 8)

The role of principals appeared to be key in implementing a reform:

Principals' attitudes towards accountability policies also influenced teachers' experiences. Principals' emphasis on or avoidance of instructional reforms, for example, influenced teachers' acceptance or critiques.

The authors of this study also note that we need to look more directly at the role of principals:

Currently, accountability policy is the primary locus of reform. While our research shows that district-level policies, have the power to facilitate or hinder implementation, the state must recognize district-level accountability policy.

The message is clear: There are many barriers to the implementation of educational policy initiatives. Policies that incorporate procedures and processes that are not mandated changes can face additional constraints.

The more fundamental issue, however, is how teachers will respond to results produced by the implementation process is not closely monitored. In medicine is not encouraging. Consider the numbers of hospital patients who are infected by the hands of doctors and/or nurses. Resnik (2004) and reporting by Atul Gawande (2004) "lines" (such as a line for a catheter) frequently lead to infection. In intensive care units, more than ten days have an infection rate of 5% to 28%, depending on how long the patient is infected. Those who survive line infections often die. Thus such infections clearly are a serious problem. More days in the hospital, more dollars in costs, more missed work.

Johns Hopkins researcher and physician designed a policy intervention focused on

conclusion running through empirical research is that it is incredibly hard to make something happen primarily outside what matters. We have learned primarily on two broad factors: local context and the attitudes, motivation, and beliefs of those who respond to a policy's goals or strategies, or implementation.

Local context plays a crucial role in implementation. Yet randomized controlled trials have failed to come to terms with the fact that they cannot be manipulated experimentally, and thus a randomized controlled trial, may be precisely what is needed in the grand scheme of schooling and education.

Public policy also suggests that local context matters. For example, in the randomized trial of the Implementation Program researchers were unable to implement the program. Instead, the reform varied to fit local contexts in Chicago. Cook explains:

The success of [Comer's] SDP [School Development Program] depends on both district and school in order to succeed. The Chicago program is implemented in the middle of a citywide effort to reform education. Local politicians, businesses and philanthropists emphasized local school councils; principals emphasized their own school goals and established their own school goals and established their own school goals and established their own school goals . . . [although this pattern] . . . did not occur.

(Cook et al. 2000: 541)

When the validity of an assessment is compromised, then the validity of an assessment-based trial is compromised. Principals play a critical role in policy implementation in California reported:

Principals have often played a pivotal role in how policies are implemented. Principals either acted as a buffer against test-score pressures, or as an added pressure emphasizing the need to raise test scores.

(PACE 2006: 8)

The role of principals appeared to be key in understanding how teachers experienced a reform:

Principals' attitudes towards accountability and their leadership styles also influenced teachers' experiences with district and state mandates. Principals' emphasis on or avoidance of certain district and state instructional reforms, for example, were reflected in teachers' acceptance or critiques.

(PACE 2006: 8)

The authors of this study also note that rather than simply looking at schools, we need to look more directly at the mediating role of districts:

Currently, accountability policy assumes that individual schools are the primary locus of reform. While school-based reform is essential, our research shows that districts, through various programs and policies, have the power to facilitate or hinder school improvement. The state must recognize district influence in the implementation of accountability policy.

(PACE 2006: 9)

The message is clear: There are many mediating factors that shape the implementation of educational policy. This is why it is crucial that reform initiatives incorporate procedures and "feedback loops" so that during the implementation, mandated changes can be modified to accommodate institutional constraints.

The more fundamental issue, however, is the degree to which institutions will respond to results produced by randomized controlled trials if the implementation process is not closely monitored. Here too, the evidence from medicine is not encouraging. Consider, for example, the surprisingly large numbers of hospital patients who are exposed to an infection from germs on the hands of doctors and/or nurses. Research by Sean Berenholtz and colleagues (2004) and reporting by Atul Gawande (2007) indicate that contamination of "lines" (such as a line for a catheter) from the germs on caregivers' hands can lead to infection. In intensive care units, patients who have a line in for more than ten days have an infection rate of 4%. Fatalities from line infections range from 5% to 28%, depending on how sick patients are when they become infected. Those who survive line infections spend more days in intensive care. Thus such infections clearly are a serious problem. They result in more deaths, more days in the hospital, more dollars spent on medical costs, and more days of missed work.

Johns Hopkins researcher and physician Peter Provonost and his colleagues designed a policy intervention focused on the steps that health practitioners

take when putting in a line: they wash their hands with soap, clean the patient's skin with an antiseptic, put sterile drapes over the patient's entire body, wear a sterile mask, hat, gown, and gloves, and put a sterile dressing over the catheter site once the line is in. The researchers discovered that about one-third of the time, the clinicians skipped one of these essential steps (Berenholtz et al. 2004; Gawande 2007). Eventually, Dr. Pronovost's team developed a checklist for a new protocol that empowered nurses to stop doctors to question them if they missed a step. Infection rates plummeted. Compared to an earlier period, the ten-day line infection rate went from 11% to 0%. In a 15-month period, there were only two line infections. Thus, adhering to the new protocol probably prevented 43 infections and eight deaths, and saved \$2 million in costs (Berenholtz et al. 2004).⁶

Yet when hospital administrators attempted to "scale up" this new model and implement it widely, the policy intervention was resisted rather than embraced (Berenholtz and Pronovost 2003; Gawande 2007). This happened, in part, because, as in schools, the intervention required a change in organizational processes. The hierarchy of authority in medicine has been weakened due to changes in insurance and managed care. Nevertheless, doctors have considerable power; nurses have lower professional status and autonomy compared to doctors. The checklist required that nurses or technicians stop the procedure (in hectic, often understaffed intensive care units) to ask, "Doctor, have you washed your hands?" This reversal of the usual hierarchy of authority was difficult to implement. In addition, many hospitals and clinics have been subject to severe budget cuts and staffing reductions, both of which make routine procedures more rushed and thus undermine attempts to add even a simple new safeguard to the standard repertoire.

How long it will be before the average doctor or nurse is apt to routinely use a checklist? Peter Pronovost is not optimistic: "We could get I.C.U. checklists in use throughout the United States in two years, if the country wanted it, [but] . . . [A]t the current rate, it will never happen" (quoted in Gawande 2007: 8). The resistance is rooted not in science or technology but in social and institutional factors:

If someone found a new drug that could wipe out infections with anything remotely like the effectiveness of Pronovost's lists, there would be television ads with Robert Jarvik extolling its virtues, detail men offering free lunches to get doctors to make it part of their practice, government programs to research it, and competitors jumping in to make a newer, better version. That's what happened when manufacturers marketed central-line catheters coated with silver or other antimicrobials: they cost a third more, and reduced infections slightly—and hospital[s] have spent tens of millions of dollars on them. But with a checklist, what we have is Peter Pronovost trying to

see if maybe, in the next year or two, New Jersey will give his idea a try.

It is rare that a policy intervention overcomes the resistance to its implementation in education for educational researchers. It is hard to get findings that are not enough; teachers, principals, and parents to be "brought along," to be helped to see that another approach is naive.

Concluding Thoughts

There are variations in how strident the push is for randomized controlled trials. But in Chapter 1, the use of this method was the headway. Increasingly, funding priorities privilege random assignment and academic career assessments over the "quality" of education research. The trend would be less alarming if decisionmakers acknowledged that the current approach to education, and that different questions are being asked from a different perspective,⁷ randomized controlled trials are not the best approach among many in the repertoire. It also would be a less troubling move if randomized controlled trials offered a more clear-eyed view of the challenges with carrying out randomized controlled trials. The combination of a narrow vision of the value of RCTs as a "gold standard" and a reference to RCTs as a "gold standard" is inferior. That such sweeping claims about the value of RCTs in education, in structurally limited opportunities schools have to do research, is of great concern. The conditions of education research are sometimes dramatically.

It is also troubling that advocates of RCTs in education, even if conditions could be made more favorable, that even if methodology can accommodate is not the best approach in education research, as in medicine, are not making the case for a different approach. Key factors in education research that does not make them any less important are the processes of educational delivery systems and the processes of a random assignment approach do not

their hands with soap, clean the patient's legs over the patient's entire body, wear a cap and put a sterile dressing over the catheter site. They discovered that about one-third of the essential steps (Berenholtz et al. 2004; Pronovost's team developed a checklist for a study to stop doctors to question them if they were not met. Compared to an earlier period, the infection rate fell from 11% to 0%. In a 15-month period, the hospital, by adhering to the new protocol, prevented 10 deaths, and saved \$2 million in costs.

When they attempted to "scale up" this new model of infection prevention, the intervention was resisted rather than embraced (Gawande 2007). This happened, in part because the change in organizational structure required a change in organizational culture. In medicine, the culture has been weakened due to the loss of status and autonomy compared to nurses or technicians stop the procedure (in intensive care units) to ask, "Doctor, have you finished the usual hierarchy of authority was broken by hospitals and clinics have been subject to external influences, both of which make routine infection prevention attempts to add even a simple checklist.

The doctor or nurse is apt to routinely use a checklist. "We could get I.C.U. checklists in two years, if the country wanted it, [but] it won't happen" (quoted in Gawande 2007: 8). The technology but in social and institu-

It could wipe out infections with the effectiveness of Pronovost's lists, there is a lot of talk extolling its virtues, detail doctors to make it part of their research it, and competitors offer a version. That's what happened with central-line catheters coated with silver. The infection rate fell a third more, and reduced infections cost tens of millions of dollars on. It is what Peter Pronovost trying to

see if maybe, in the next year or two, hospitals in Rhode Island and New Jersey will give his idea a try.

(Gawande 2007: 7)

It is rare that a policy intervention is as clearly effective as the checklist. Thus the resistance to its implementation is a particularly important cautionary tale for educational researchers. It is hard to implement research results. Valid findings are not enough; teachers, parents, students, and administrators need to be "brought along," to be helped to "buy into" the proposed change. Any other approach is naive.

Concluding Thoughts

There are variations in how strident researchers are in insisting on the need for randomized controlled trials. But, as Pamela Barnhouse Walters has noted in Chapter 1, the use of this methodology in education research has made headway. Increasingly, funding priorities for research grants and for professional training resources privilege random assignment; journal reviews, grant reviews, and academic career assessments consider this criterion in evaluating the "quality" of education research. The term "gold standard" is common.

This trend would be less alarming if federal department of education decisionmakers acknowledged that there are many varied research questions in education, and that different questions call for different methods. From this perspective, randomized controlled trials would be one, and only one, approach among many in the repertoire available to social science researchers. It also would be a less troubling movement if advocates of randomized controlled trials offered a more clear-eyed assessment of the problems associated with carrying out randomized controlled trials. Thus what is problematic is the combination of a narrow vision and the belief, signaled by the frequent reference to RCTs as a "gold standard," that all other methods are, in crucial ways, inferior. That such sweeping claims are made despite the flawed and structurally limited opportunities schools offer for carrying out randomized controlled trials is of great concern. Educational institutions are not ideal places to do research. The conditions of the "laboratory" change frequently, and sometimes dramatically.

It is also troubling that advocates of RCTs do not sufficiently acknowledge that, even if conditions could be made ideal, the range of questions this methodology can accommodate is narrow. Many important questions in education research, as in medicine, are not amenable to a random assignment approach. Key factors in education research are difficult to manipulate, but this does not make them any less important. Studying the meaning and social processes of educational delivery systems is crucial, and it is an area advocates of a random assignment approach do not sufficiently acknowledge.

The current movement toward the use of randomized controlled trials in education research fails to offer an umbrella sufficiently broad to both welcome and respect diverse methodological approaches. This narrowness of spirit is not only a tactical mistake. Its short-sightedness will, in the end, impede random assignment advocates from realizing the educational reform goals they hold most dear.

Notes

- * The author is grateful to Dana Burke, Aaron Pallas, Pamela Barnhouse Walters, and the anonymous reviewers for helpful comments on an earlier version of this chapter. All errors, of course, are the responsibility of the author.
1. For example, at the top of the home page of the U.S. Department of Education, Institute of Education Sciences, National Center for Education Research is the statement: "The National Center for Education Research (NCER) supports rigorous research that addresses the nation's most pressing education needs, from early childhood to adult education." Under "What's New," an announcement for a summer research institute on cluster-randomized trials explains: "The purpose of the Summer Research Training Institute on Cluster-Randomized Trials is to increase the national capacity of researchers to develop and conduct rigorous evaluations of the effectiveness of education interventions." Retrieved December 18, 2007 from <http://www.nces.ed.gov> and <http://www.ed.gov/whatsnew/conferences?pid=310&cid=3>. See also U.S. Department of Education (2003a, 2003b) and Lewis (2003).
2. Given the evidence that exists about the effectiveness of policy interventions of the last 40 years, it is foolhardy to search for a single policy intervention or "silver bullet" to improve the educational achievement for all American children in all schools all of the time. There are also likely to be interactions. Doris Entwisle and Karl Alexander at the Johns Hopkins University, for example, showed that teachers who themselves had low socioeconomic origins did particularly well in producing achievement from students with low socioeconomic origins; these teachers from low socioeconomic origins outperformed teachers who had been raised in more prosperous families (Entwisle, Alexander, and Olson 1997). Thus the impact of policies that are universally implemented is not universal in the student population. Some students benefit more than others.
3. In the development of clinical trials for drugs, neither the physician nor the patient knew if the patient was in the subgroup that was receiving a "placebo" or "sugar pill" or if he/she was actually receiving the treatment, hence the term "double-blind." In this research tradition, every effort is made to select individuals who are so similar to one another that any difference in their outcomes could be attributed to the treatment. It is this approach that has been adopted for schools.
4. Compliance is higher for acute illnesses than for chronic conditions. In clinical trials, patients tend to be more dutiful due to the follow-up, but even under these conditions medication adherence rates are reported to be only in the range of 43% to 78% (see Osterberg and Blaschke 2005).
5. Some teachers are more effective in promoting parent involvement than others; see Becker and Epstein's work (1982) on teacher leaders.
6. Pronovost and colleagues stress the importance of retaining professional expertise rather than dictating a complex set of procedures. They explain: "We reduced our rate of [infection] using relatively simple and inexpensive interventions, as opposed to implementing more expensive interventions, such as antibiotic/antiseptic catheters. For interventions to work in the busy world of clinical practice, they should be simple to implement. By changing systems rather than exhorting providers to comply with guidelines, we can help ensure that patients receive effective therapies. For example, it was difficult to write a detailed guideline regarding the need for a central venous catheter; there are too many decisions to account for. It is unlikely that detailed guidelines would be practical for complex decisions, such as ICU admission and discharge, extubation, and use of catheters. Rather, we simply asked physicians to consider daily whether central catheters could be removed, highlighting the risk of catheters yet allowing physicians to use their clinical judgment" (Berenholtz et al. 2004: 2020).
7. A perspective that most first-year graduate students in the social sciences and in education encounter in their beginning research methods courses, by the way: that the trick is to match the research method to the question the researcher wants to investigate, not to find one one-size-fits-all "best" method.

References and Further Reading

- Anyon, Jean. 2005. *Radical Possibilities: Public Schools in New York*. Teachers College Press.
- Becker, H. and J. Epstein. 1982. "Parent Involvement in Schools." *School Journal* 83(2):85-102.
- Berenholtz, Sean M., Peter J. Pronovost, Pamela E. Farley, Shelley Milanovich, Elizabeth G. Dorman, and Trish M. Perl. 2004. "Eliminating Intensive Care Unit." *Critical Care Medicine* 32(12):2145-2150.
- Berenholtz, Sean M. and Peter J. Pronovost. 2008. "Current Opinion Critical Care 9:521-525."
- Burton, Linda M., Townsend Price-Spratlen, and about Measuring Neighborhoods: Implications for Children." Pp. 132-144 in *Measuring Neighborhoods*, edited by Jeanne Brooks-Gunn and Russell Sage.
- Chronicle of Higher Education*. 2008. "Education Research: A New Focus." March 7, 2008. <http://chronicle.com/wire/2008/03/07>.
- Clampet-Lundquist, Susan, Kathryn Edin, Jeffrey J. Reitzel, and Susan M. Edin. 2007. "Risk Youth out of High-Risk Neighborhoods: A Policy Paper, Industrial Relations Section, Princeton University." <http://www.irsc.princeton.edu/papers/20070701.pdf>.
- Comer, James P. 1995. *School Power*. New York: Basic Books.
- Comer, James P. and Norris M. Haynes. 1999. "The 'Comer' School Development Program in Chicago: An Evaluation" by Thomas D. Cook et al. *Journal of Research in Education* 20(1):1-12.
- Conrad, Clifton F. and Ronald C. Serlin (eds.). 2002. *Engaging Ideas and Enriching Inquiry*. Thousand Oaks, CA: Sage.
- Cook, Thomas D. and Monique Payne. 2002. "Assignment in Educational Research." Pp. 1-12 in *Educational Research*, edited by F. Mosteller and D. T. S. Moore.
- Cook, Thomas D., Robert F. Murphy, and H. F. O'Neil. 2002. "Program in Chicago: A Theory-Based Evaluation." *Journal of Research in Education* 23(2):535-597.
- Cook, Thomas D., Farah-Naaz Habib, Meredith P. Serdar M. Degirmencioglu. 1998. "Comer's School Development Program: A Theory-Based Evaluation." *Journal of Research in Education* 19(3):543-597.
- DeParle, Jason. 2004. *American Dream: Three Years in New York*. New York: Viking.
- Diamond, John B. and Kimberley Gomez. 2004. "Schools: The Implications of Social Class on Urban Society." *Journal of Research in Education* 25(3):383-427.
- Duncan, Greg J., Aletha C. Huston, and Thomas J. Kane. 2007. *Working Poor and Their Families*. New York: Russell Sage.
- Edin, Kathryn and Laura Lein. 1997. *Making Ends Meet: How Working Poor Families Survive*. New York: Russell Sage.
- Educational Researcher*. 2002. Theme issue on Schools. Retrieved December 18, 2007 from <http://er.sagepub.com>.
- Entwisle, Doris R. and Karl L. Alexander. 1992. "The Role of Schools in the Life of the Inner-City Child." *Journal of Research in Education* 13(1):77-84.
- Entwisle, Doris R., Karl L. Alexander, and Linda L. Olson. 1997. *The Life of the Inner-City Child*. Boulder, CO: Westview Press.
- Epstein, Joyce Levy. 2002. *School, Family, and Community Partnerships: Making It Work*. New York: Corwin Press.
- Ericksen, Kai. 1978. *Everything in Its Path*. New York: Basic Books.
- Erickson, Frederick and Kris Gutierrez. 2002. "Schools and Communities: A New Paradigm." *Journal of Research in Education* 23(1):21-24.
- Furman, Jason and Robert E. Rubin. 2007. *Using Economic Imperative*. Washington, DC: Urban Institute.

use of randomized controlled trials in health care is not sufficiently broad to both welcome and challenge such approaches. This narrowness of spirit is not surprising, in the end, because the educational reform goals they hold

James P. Pallas, Pamela Barnhouse Walters, and the author on an earlier version of this chapter. All errors, of course, are my own.

of the U.S. Department of Education, Institute of Education Research is the statement: "The National Longitudinal Study of the Youth reports rigorous research that addresses the nation's challenge to adult education." Under "What's New," the Institute on Cluster-Randomized Trials explains, "The Institute on Cluster-Randomized Trials is to increase the use of cluster-randomized trials to increase the development and conduct of rigorous evaluations of the effectiveness of policy interventions of the last 40 years."

of intervention or "silver bullet" to improve the health of all schools all of the time. There are also Karl Alexander at the Johns Hopkins University, for example, who had low socioeconomic origins did particularly well on tests with low socioeconomic origins; these teachers of teachers who had been raised in more prosperous areas. Thus the impact of policies that are universally applied. Some students benefit more than others. For example, neither the physician nor the patient knew if the pill was a "placebo" or "sugar pill" or if he/she was actually blind. In this research tradition, every effort is made to ensure that any difference in their outcomes is due to the approach that has been adopted for schools.

for chronic conditions. In clinical trials, patients are often randomly assigned to treatment or control groups, but even under these conditions medication adherence ranges from 43% to 78% (see Osterberg and Blaschke 2006).

ing parent involvement than others; see Becker and

of retaining professional expertise rather than explain: "We reduced our rate of [infection] using antibiotics, as opposed to implementing more expensive interventions. For interventions to work in the busy world of health care, we can help ensure that patients receive effective care by changing systems rather than blaming individuals. A detailed guideline regarding the need for a central team to account for. It is unlikely that detailed guidelines for ICU admission and discharge, extubation, and weicians to consider daily whether central catheters are necessary yet allowing physicians to use their clinical

students in the social sciences and in education is courses, by the way: that the trick is to match the research to what you want to investigate, not to find one one-size-fits-

References and Further Reading

- Anyon, Jean. 2005. *Radical Possibilities: Public Policy, Urban Education, and a New Social Movement*. New York: Teachers College Press.
- Becker, H. and J. Epstein. 1982. "Parent Involvement: A Survey of Teacher Practices." *The Elementary School Journal* 83(2):85-102.
- Berenholtz, Sean M., Peter J. Pronovost, Pamela A. Lipsett, Deborah Hobson, Karen Earsing, Jason E. Farley, Shelley Milanovich, Elizabeth Garrett-Mayer, Bradford Winters, Haya R. Rubin, Todd Dorman, and Trish M. Perl. 2004. "Eliminating Catheter-Related Bloodstream Infections in the Intensive Care Unit." *Critical Care Medicine* 32:2014-2020.
- Berenholtz, Sean M. and Peter J. Pronovost. 2003. "Barriers to Translating Evidence into Practice." *Current Opinion Critical Care* 9:321-325.
- Burton, Linda M., Townsend Price-Spratlen, and Margaret Beale Spencer. 1997. "On Ways of Thinking about Measuring Neighborhoods: Implications for Studying Context and Developmental Outcomes for Children." Pp. 132-144 in *Neighborhood Poverty: Policy Implications in Studying Neighborhoods*, edited by Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber. New York: Russell Sage.
- Chronicle of Higher Education. 2008. "Education Dept. to End Controversial Study of Upward Bound." March 7, 2008. <http://chronicle.com/weekly/v54/26/26a02001.htm>. Accessed September 22, 2008.
- Clampet-Lundquist, Susan, Kathryn Edin, Jeffrey B. Kling, and Greg J. Duncan. 2006. "Moving At-Risk Youth out of High-Risk Neighborhoods: Why Do Girls Fare Better than Boys?" Working Paper, Industrial Relations Section, Princeton University.
- Cornier, James P. 1995. *School Power*. New York: Free Press.
- Cornier, James P. and Norris M. Haynes. 1999. "The Dynamics of School Change: Response to the Article 'Cornier's School Development Program in Prince George's County, Maryland: A Theory-Based Evaluation' by Thomas D. Cook et al." *American Educational Research Journal* 36(3):599-607.
- Conrad, Clifton F. and Ronald C. Serlin (eds.) 2005. *The SAGE Handbook for Research in Education: Engaging Ideas and Enriching Inquiry*. Thousand Oaks, CA: Sage Publications.
- Cook, Thomas D. and Monique Payne. 2002. "Objecting to the Objections to Using Randomized Assignment in Educational Research." Pp. 150-179 in *Evidence Matters: Randomized Trials in Education Research*, edited by F. Mosteller and R. Boruch. Washington, DC: Brookings Institution Press.
- Cook, Thomas D., Robert F. Murphy, and H. David Hunt. 2000. "Cornier's School Development Program in Chicago: A Theory-Based Evaluation." *American Educational Research Journal* 37(2):533-597.
- Cook, Thomas D., Farah-Naaz Habib, Meredith Phillips, Richard A. Settersten, Shobha C. Shagle, and Serdar M. Degirmencioglu. 1999. "Cornier's School Development Program in Prince George's County, Maryland: A Theory-Based Evaluation." *American Educational Research Journal* 36(3):543-597.
- DeParle, Jason. 2004. *American Dream: Three Women, Ten Kids, and a Nation's Drive to End Welfare*. New York: Viking.
- Diamond, John B. and Kimberley Gomez. 2004. "African-American Parents' Orientations towards Schools: The Implications of Social Class and Parents' Perceptions of Schools." *Education and Urban Society* 36:383-427.
- Duncan, Greg J., Aletha C. Huston, and Thomas S. Weisner. 2007. *Higher Ground: New Hope for the Working Poor and Their Families*. New York: Russell Sage Foundation.
- Edin, Kathryn and Laura Lein. 1997. *Making Ends Meet: How Single Mothers Survive Welfare and Low-Wage Work*. New York: Russell Sage.
- Educational Researcher. 2002. Theme Issue on *Scientific Research and Education*. 31:3-29.
- Entwisle, Doris R. and Karl L. Alexander. 1992. "Summer Setback." *American Sociological Review* 57:72-84.
- Entwisle, Doris R., Karl L. Alexander, and Linda Steffel Olson. 1997. *Children, Schools, and Inequality*. Boulder, CO: Westview Press.
- Epstein, Joyce Levy. 2002. *School, Family, and Community Partnership: Your Handbook for Action*. New York: Corwin Press.
- Erickson, Kai. 1978. *Everything in Its Path*. New York: Simon and Schuster.
- Erickson, Frederick and Kris Gutierrez. 2002. "Science' Rejects Postmodernism." *Educational Researcher* 31(8):21-24.
- Furman, Jason and Robert E. Rubin. 2007. *Universal, Effective, and Affordable Health Insurance: An Economic Imperative*. Washington, DC: Brookings Institution Press.

- Gawande, Atul. 2007. "Annals of Medicine: The Checklist." *The New Yorker*. December 21, 2007. Retrieved January 2, 2008 (http://www.newyorker.com/reporting/2007/12/10/071210fa_fac_gawande?printable=true).
- Hallinan, Maureen T. (ed.). 2006. *Handbook of the Sociology of Education*. New York: Springer.
- Haynes, R. Brian, Heather B. McDonald, and Amit X. Garg. 2002. "Helping Patients Follow Prescribed Treatment." *Journal of the American Medical Association (JAMA)* 288:2880-2883. Retrieved December 11, 2002 (<http://www.jama.com>).
- Howe, Kenneth and Margaret Eisenhart. 1990. "Standards for Qualitative (and Quantitative) Research: A Prolegomenon." *Educational Researcher* 19(4):2-9.
- Hubbard, Lea and Hugh Mehan. 1999a. "Scaling Up an Un-tracking Program: A Co-constructed Process." *Journal of Education for Students Placed at Risk* 4(1):83-100.
- Hubbard, Lea and Hugh Mehan. 1999b. "Race and Reform: Educational 'Niche Picking' in a Hostile Environment." *Journal of Negro Education* 68(2):213-226.
- Hubbard, Lea, Mary Kay Stein, and Hugh Mehan. 2006. *Reform as Learning: When School Reform Collides with School Culture and Community Politics*. London: Routledge.
- Iceland, John. 2005. "Measuring Poverty: Theoretical and Empirical Considerations." *Measurement* 3(4):207-243.
- JESPAR (*Journal of Education for Students Placed at Risk*). 1998. "Changing Schools for Changing Times: The Corner School Development Program." Special issue of the *Journal of Education for Students Placed at Risk* (JESPAR). 3(1):3-19.
- Kaestle, Carl F. 1993. "The Awful Reputation of Education Research." *Educational Researcher* 22(1):23-26-31.
- Kahne, Joseph and Kim Bailey. 1999. "The Role of Social Capital in Youth Development: The Case of 'I Have a Dream' Programs." *Educational Evaluation and Policy Analysis* 21(3):321-343.
- Karabel, Jerome and A. H. Halsey (eds.). 1977. *Power and Ideology in Education*. Oxford: Oxford University Press.
- Klimberg, Eric. 2002. *Everything in Its Path*. Chicago: University of Chicago Press.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75:83-119.
- Kozol, Jonathan. 1991. *Savage Inequalities: Children in America's Schools*. New York: Crown.
- Labaree, David F. and Aaron M. Pallas. 1996. "Dire Straits: The Narrow Vision of the Holmes Group." *Educational Researcher* 25(4):25-28.
- Lamont, Michele. 2005. "Criteria for Evaluation of Qualitative Research." Workshop on Interdisciplinary Standards for Systematic Qualitative Research. National Science Foundation Cultural Anthropology, Political Science, Sociology, and Law and Social Science Programs. May 19-20, 2005. Washington, DC. Retrieved December 15, 2007 (<http://www.wjh.harvard.edu/nsfqual/>).
- Lareau, Annette. 2000. *Home Advantage: Social Class and Parental Intervention in Elementary Education*. Updated edition. Lanham, MD: Rowman and Littlefield.
- . 2003. *Unequal Childhoods: Class, Race, and Family Life*. Berkeley, CA: University of California Press.
- Lareau, Annette and Jeffrey Shultz (eds.). 1999. *Journeys through Ethnography: Realistic Accounts of Fieldwork*. Boulder, CO: Westview.
- Lewis, Anne C. 2003. "New Hope for Educational Research?" *Phi Delta Kappan* 84(5):339-340.
- Lowes, Robert. 1998. "Patient-Centered Care for Better Patient Adherence." *Family Practice Management* 5(3). Retrieved December 10, 2007 (<http://www.aafp.org/lpm/980306fm/patient.html>).
- McLaughlin, Mäbrey Wallin. 1987. "Learning from Experience: Lessons from Policy Implementation." *Educational Evaluation and Policy Analysis* 9(2):171-178.
- Montori, Victor M., Yaqian Grace Wang, Pablo Alonso-Coello, and Samit Bhagra. 2006. "Systematic Evaluation of the Quality of Randomized Controlled Trials in Diabetes." *Diabetic Care* 29:1833-1838.
- Mosteller, Frederick and Robert Boruch (eds.). 2002. *Evidence Matters: Randomized Trials in Education Research*. Washington, DC: Brookings Institution Press.
- National Council on Patient Information and Education. 2007. *Enhancing Prescription Medicine Adherence: A National Action Plan*. Bethesda, MD: National Council on Patient Information and Education.
- Osterberg, Lars and Terrence Blaschke. 2005. "Adherence to Medication: Review Article." *New England Journal of Medicine* 353:5. Retrieved September 22, 2008 (<http://www.nejm.org>).
- PACE (Policy Analysis for California Education). 2006. *Voices from the Field: Educators Respond to Accountability*. Elisabeth Woody, Melissa Buttes, Judith Kafka, Sandra Park, and Jennifer Russell. University of California, Berkeley. Retrieved September 22, 2008 (http://pace.berkeley.edu/ERAP_Report-WEB.pdf).
- Pierson, Paul. 2004. *Politics in Time: History and the Politics of the Present*. Princeton, NJ: Princeton University Press.
- Ragin, Charles, Ioane Nagel, and Patricia V. Ray. 2007. "Comparative Perspective." *Journal of Education for Students Placed at Risk* 12(1):1-10. Retrieved December 2, 2007 (<http://www.naf.org>).
- Rainwater, Lee and Timothy Smeeding. 2005. *Comparative Perspective*. New York: Routledge.
- Reich, Jennifer A. 2005. *Fixing Families: How Schools Can Help*. New York: Routledge.
- Research for Action. 2005. "Children and Education Program." Philadelphia: www.researchforaction.org.
- Rumberger, Russell W. and Scott L. Thomas. 2005. "The Achievement Gap: A Comparative Perspective." *Journal of Education for Students Placed at Risk* 10(1):1-10.
- Ryan, Gery W. 2005. "What Are Standard Interdisciplinary Standards for Social Cultural Anthropology, Political Science, and Sociology?" Washington, DC: Research for Action.
- Sunderman, Gail L., James S. Kim, and Gary A. Cook. 2004. *From the Field*. Thousand Oaks, CA: Corwin.
- Swanson, Christopher B. 2004. "The Real Story." *Education Policy Center*. Retrieved September 22, 2008 (<http://www.cep.org>).
- Teachers College Record. 2005. A Symposium Report for Qualitative Inquiry. 107:1-107:10.
- Tierney, Joseph P., Jean Baldwin Grossman, and Joseph P. Grossman. 2007. *Study of Big Brothers/Big Sisters*. Philadelphia, PA: PPV.org.
- U.S. Department of Education, Institute of Education Sciences, Office of Educational Evaluation and Regional Assistance. *Supported by Rigorous Evidence: A Report for the U.S. Department of Education*. Washington, DC: Printing Office.
- U.S. Department of Education, Office of the Inspector General. *Behind a Parents Guide*. Washington, DC: Printing Office.
- U.S. Department of Education, Institute of Education Sciences, Office of Educational Evaluation and Regional Assistance. *Trusted Source of Scientific Evidence*. Washington, DC: Printing Office.
- U.S. Department of Education. 2007. *The Condition of Education 2007*. Washington, DC: Printing Office.
- Whitehurst, Grover "Russ." 2002. "Educational Achievement and Student Accountability." Washington, DC: Institute for Education Sciences, Washington, DC.

Checklist." *The New Yorker*, December 21, 2007. www.newyorker.com/reporting/2007/12/10/071210fa_fac.

in *Sociology of Education*. New York: Springer.

in X. Garg. 2002. "Helping Patients Follow Prescribed Risk Association (JAMA) 288:2880-2883. Retrieved

Standards for Qualitative (and Quantitative) Research: 1941-2-6.

ing Up an Un-tracking Program: A Co-constructed *Plan of Risk* 4(1):83-100.

and Reform: Educational 'Niche Picking' in a Hostile 86(2):213-226.

in. 2006. *Reform as Learning: When School Reform by Politics*. London: Routledge.

tical and Empirical Considerations." *Measurement*

in Risk. 1998. "Changing Schools for Changing Times: Special Issue of the *Journal of Education for Students*

Education Research." *Educational Researcher* 22(1):23.

of Social Capital in Youth Development: The Case of *Evaluation and Policy Analysis* 21(3):321-343.

Power and Ideology in Education. Oxford: Oxford

University of Chicago Press.

rence F. Katz. 2007. "Experimental Analysis of 1-119.

in *America's Schools*. New York: Crown.

in *Strait: The Narrow Vision of the Holmes Group*."

Qualitative Research." Workshop on Interdisciplinary research. National Science Foundation Cultural and Law and Social Science Programs. May 19-20, 2007 (<http://www.wjh.harvard.edu/nsfqual/>).

in *and Parental Intervention in Elementary Education*, and Littlefield.

Family Life. Berkeley, CA: University of California

Journeys through Ethnography: Realistic Accounts of Research? *Phi Delta Kappan* 84(5):339-340.

for Better Patient Adherence." *Family Practice* 2007 (<http://www.aafp.org/fpm/980300fm/patient>).

in Experience: Lessons from Policy Implementation." 9(2):171-178.

Alonso-Coello, and Sumit Bhagra 2006. "Systematic and Controlled Trials in Diabetes." *Diabetic Care*

62. *Evidence Matters: Randomized Trials in Education* itation Press.

Education. 2007. *Enhancing Prescription Medicine* a, MD: National Council on Patient Information and

herence to Medication: Review Article." *New England* iber 22, 2008 (<http://www.nejm.org>).

il. 2006. *Voices from the Field: Educators Respond to* ities. Judith Kafka, Sandra Park, and Jennifer Russell. ed September 22, 2008 (<http://pacc.berkeley.edu/>

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.

Ragin, Charles, Joane Nagel, and Patricia White. 2004. *Workshop on Scientific Foundations of Qualitative Research*. Sociology Program, National Science Foundation, July 11-12, 2003. Retrieved December 2, 2007 (<http://www.nsf.gov/pubs/2004/nsf04219/nsf04219.pdf>).

Rainwater, Lee and Timothy Smeeding. 2003. *Poor Kids in a Rich Country: America's Children in Comparative Perspective*. New York: Russell Sage.

Reich, Jennifer A. 2005. *Fixing Families: Parents, Power, and the Child Welfare System*. New York: Routledge.

Research for Action. 2005. "Children and Families First: An Evaluation of the Philadelphia Say Yes to Education Program." Philadelphia, PA, November. Retrieved November 20, 2007 (<http://www.researchforaction.org>).

Rumberger, Russell W. and Scott L. Thomas. 2000. "The Distribution of Dropout and Turnover Rates among Urban and Suburban High Schools." *Sociology of Education* 73(1):39-67.

Ryan, Gery W. 2005. "What Are Standards of Rigor for Qualitative Research?" Workshop on Interdisciplinary Standards for Systematic Qualitative Research. National Science Foundation Cultural Anthropology, Political Science, Sociology, and Law and Social Science Programs. May 19-20, 2005. Washington, DC. Retrieved September 22, 2008 (<http://www.wjh.harvard.edu/nsfqual/>).

Sunderman, Gail L., James S. Kim, and Gary Orfield. 2005. *NCLB Meets School Realities: Lessons from the Field*. Thousand Oaks, CA: Corwin Press.

Swanson, Christopher B. 2004. "The Real Truth about Low Graduation Rates: An Evidence-Based Commentary." Education Policy Center, The Urban Institute, Washington, DC. Retrieved September 22, 2008 (http://www.urban.org/UploadedPDF/411050_realtruth.pdf).

Teachers College Record. 2005. A Symposium on the Implications of the Scientific Research in Education Report for Qualitative Inquiry. 107:1-58.

Tierney, Joseph P., Jean Baldwin Grossman, with Nancy L. Resch. 2000. *Making a Difference: An Impact Study of Big Brothers/Big Sisters*. Philadelphia, PA: Public/Private Ventures. Retrieved November 28, 2007 (<http://www.ppv.org>).

U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. 2003a. *Identifying and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, Office of the Secretary, Office of Public Affairs. 2003b. *No Child Left Behind: A Parents Guide*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, Institute of Education Sciences. 2006. *What Works Clearinghouse: A Trusted Source of Scientific Evidence of What Works in Education*. Washington, DC. Retrieved December 12, 2007 (<http://www.whatworks.ed.gov/>).

U.S. Department of Education. 2007. *The Condition of Education*. Washington, DC: U.S. Government Printing Office.

Whitehurst, Grover "Russ." 2002. "Evidence-Based Education." Presentation at the Student Achievement and Student Accountability Conference. U.S. Department of Education, Institute for Education Sciences. Washington, DC, October 8, 2002.