Educational research is not what it was fifteen years ago. In this discipline the changes have been dramatic, far-reaching and rapid. Out of the criticisms of the 1990s and the calls for greater accountability of the early 2000s one idea has come to the fore—education science. There are two main components to education science. The first is the principle that research in education must model itself more closely on disciplines seen as highly credible and successful, mainly the natural sciences and medicine. The second part is that educators must build their practice upon the insights developed through this scientific research process. Overall, education science has the potential to change how we think about education, how we build knowledge about it, and how we know when it is successful.

This volume brings together some of the most active proponents of education science and some of the most committed critics. Within it the idea of education science is explored in depth, randomized controlled trials (considered the “gold standard” of education science) are discussed in detail, and the philosophical difficulties of knowledge in education are explored. Established thinkers are brought alongside newly emerging analysts, and detailed accounts of the institutions driving education science are included. Each contribution is thoughtful and balanced, engaging with the issues of the field and how they might be addressed. As a body of work, this collection of essays provides a well-rounded, critical discussion of the potential—and the problems—of the education science movement.
Scope

Cultural studies provides an analytical toolbox for both making sense of educational practice and extending the insights of educational professionals into their labors. In this context Transgressions: Cultural Studies and Education provides a collection of books in the domain that specify this assertion. Crafted for an audience of teachers, teacher educators, scholars and students of cultural studies and others interested in cultural studies and pedagogy, the series documents both the possibilities of and the controversies surrounding the intersection of cultural studies and education. The editors and the authors of this series do not assume that the interaction of cultural studies and education devalues other types of knowledge and analytical forms. Rather the intersection of these knowledge disciplines offers a rejuvenating, optimistic, and positive perspective on education and educational institutions. Some might describe its contribution as democratic, emancipatory, and transformative. The editors and authors maintain that cultural studies helps free educators from sterile, monolithic analyses that have for too long undermined efforts to think of educational practices by providing other words, new languages, and fresh metaphors. Operating in an interdisciplinary cosmos, Transgressions: Cultural Studies and Education is dedicated to exploring the ways cultural studies enhances the study and practice of education. With this in mind the series focuses in a non-exclusive way on popular culture as well as other dimensions of cultural studies including social theory, social justice and positionality, cultural dimensions of technological innovation, new media and media literacy, new forms of oppression emerging in an electronic hyperreality, and postcolonial global concerns. With these concerns in mind cultural studies scholars often argue that the realm of popular culture is the most powerful educational force in contemporary culture. Indeed, in the twenty-first century this pedagogical dynamic is sweeping through the entire world. Educators, they believe, must understand these emerging realities in order to gain an important voice in the pedagogical conversation.

Without an understanding of cultural pedagogy’s (education that takes place outside of formal schooling) role in the shaping of individual identity–youth identity in particular–the role educators play in the lives of their students will continue to fade. Why do so many of our students feel that life is incomprehensible and devoid of meaning? What does it mean, teachers wonder, when young people are unable to describe their moods, their affective affiliation to the society around them. Meanings provided young people by mainstream institutions often do little to help them deal with their affective complexity, their difficulty negotiating the rift between meaning and affect. School knowledge and educational expectations seem as anachronistic as a ditto machine, not that learning ways of rational thought and making sense of the world are unimportant.

But school knowledge and educational expectations often have little to offer students about making sense of the way they feel, the way their affective lives are shaped. In no way do we argue that analysis of the production of youth in an electronic mediated world demands some “touchy-feely” educational superficiality. What is needed in this context is a rigorous analysis of the interrelationship between pedagogy, popular culture, meaning making, and youth subjectivity. In an era marked by youth depression, violence, and suicide such insights become extremely important, even life saving. Pessimism about the future is the common sense of many contemporary youth with its concomitant feeling that no one can make a difference.

If affective production can be shaped to reflect these perspectives, then it can be reshaped to lay the groundwork for optimism, passionate commitment, and transformative educational and political activity. In these ways cultural studies adds a dimension to the work of education unfulfilled by any other sub-discipline. This is what Transgressions: Cultural Studies and Education seeks to produce—literature on these issues that makes a difference. It seeks to publish studies that help those who work with young people, those individuals involved in the disciplines that study children and youth, and young people themselves improve their lives in these bizarre times.
Education Science: Critical Perspectives

Ralf St. Clair
University of Glasgow, United Kingdom
# TABLE OF CONTENTS

Acknowledgements vii

Introduction ix
  
  *Ralf St. Clair*

1. Institute of Education Sciences – Putting Science Back into Education Research 1
  
  *Lynn Okagaki, Elizabeth R. Albro and Jacqueyn A. Buckley*

2. Taking Stock of Educational Research and the Impact of the UK Teaching and Learning Research Program 25
  
  *Donald Christie and Andrew Pollard*

3. An Uneasy Relationship? Teacher Education Research and Education Science 41
  
  *Ian Menter*

4. Systematic Reviews of Research on Educational Programs: Methodological and Substantive Issues 53
  
  *Robert E. Slavin*

5. Randomized Controlled Trials in Education Research 71
  
  *Carole J. Torgerson and David J. Torgerson*

6. The Future is Random – Why the RCT should often be The Method of Evaluation 83
  
  *Ben Styles*

7. What is Evidence for Evidence-Based Practice? 101
  
  *Martyn Hammersley*

8. Scientism and Scientificity in the Rage for Accountability 113
  
  *Patti Lather*

9. Meager Knowledge and Knowing Mess: “The Odor of Elephants after Rain” 127
  
  *Alison Phipps*
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10. To Infinity and Beyond: On the Social Surrounding</td>
<td>143</td>
</tr>
<tr>
<td>Science</td>
<td></td>
</tr>
<tr>
<td>Ralf St. Clair</td>
<td></td>
</tr>
<tr>
<td>Author Biographies</td>
<td>157</td>
</tr>
<tr>
<td>Index</td>
<td>161</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

Thanks go to the Faculty of Education at the University of Glasgow for supporting this work, to the many people who have been good enough to discuss these issues with me, and to the authors for their hard work and their patience.

Many thanks to Jeremiah for the idea of a maze.

To ms; thank you for the kitchen table.
RALF ST. CLAIR

INTRODUCTION

This book is an attempt to understand in more depth one of the most substantial developments in the history of educational research. In the last few years the field may have experienced the most radical attempt to change the face of a research discipline ever experienced in the social sciences: the promotion of “education science.” This term represents a complex re-alignment of factors affecting educational research so as to promote particular forms of knowledge generation and application of evidence, with many of the epistemological and methodological criteria derived from the natural and medical sciences. It is a remarkably far-reaching endeavor, affecting funding, employment and training of researchers and educators, dissemination of knowledge, and the way educational initiatives are conceived and valued.

While previous studies on education sciences have often taken one side or another in what may easily degenerate into a bitter partisan debate, it does neither side any service to caricature the other, or dismiss their concerns as missing the point. In any case, education science has achieved the status of a serious idea that researchers cannot afford to ignore. This volume seeks to understand the positions of diverse stakeholders in some depth, and support readers to draw their own conclusions based on the perspectives presented. Whatever position researchers take regarding methodology and epistemology, they are operating against a backdrop of education science. This does not imply, however, that these ideas should be taken for granted or accepted wholesale. There is still much to be done in developing and refining the principles of science appropriate to education, and how they might be applied. If there is an overarching agenda for this volume, it is this: education science is a powerful and interesting idea that needs to be approached with care and a relatively sophisticated understanding of the issues.

The history of education science, the impetus for it, and the structures supporting it, are well addressed in the chapters of this book, but it may be useful to sketch in a few lines covering the major milestones. In the UK, there is general agreement that the beginning of interest in education science can be traced to David Hargreaves’ (1996) speech to the Teacher Training Agency. In this speech, which can fairly be characterized as a thought experiment, Hargreaves asks aloud what would happen if educational research were to develop research structures similar to those in medicine along with a culture of applying the products systematically to practice. This speech was followed over the next few years by several reports quite critical of educational research (see, for example, Tooley & Darby, 1998) and the involvement of the Economics and Social Research Council in the
development of strategies to address the alleged problems. These strategies include the increasing emphasis on evidence-based practice, the Teaching and Learning Research Project (see Christie and Pollard, Chapter 2), systematic reviews of existing research, and considerable attention to research capacity building (St.Clair & Belzer, 2007).

In the US, changes were also made over the same period, primarily in response to the “No Child Left Behind Act” (NCLB, 2002). It is worth noting that even though the Act and the development of education science are strongly associated with the Bush White House, the General Accounting Office (2001) identified almost 700 points of concern with the Department of Education under Clinton, and there were already moves afoot to increase the use of evidence in education. NCLB, however, very strongly and explicitly prioritized scientific models of research on education, and created new institutions to promote them (see Okagaki, Albro & Buckley, Chapter 1). There was also considerable orientation of funding towards projects within the education science framework, provision of advice to educators on assessing evidence (with randomized controlled trials as the “gold standard”), and some efforts to ensure that the Department of Education was publicizing only “scientific” research. One of the greatest challenges was that there were so few studies meeting the retrospectively applied criteria of education science (Belzer & St.Clair, 2005).

Developments on each side of the Atlantic have several features in common. These included direction of resources towards a particular form of research, the creation of criteria and an associated hierarchy for “research quality,” the development of new institutional forms, and a recognition that both producers and consumers of educational research had to be engaged. In practice, this created a managerialist approach to educational research that fit well with the radical audit culture of the early 2000s and the re-emergence of technicism (St.Clair & Belzer, 2007).

The changes have been very rapid over the last ten years, and there is much to report. This volume sets out to look at the developments underpinned by education science, to look in some detail at the “gold standard” randomized controlled trials, and to consider some of the challenges that lie ahead for the evolution of education science. A lot of the discussion in this book captures discussions in progress, highlighting issues still to be resolved rather than questions to which there are clear answers. This includes, particularly, concerns with the boundaries of education science and its relationship with wider questions of education. More immediately, researchers and other educational thinkers are just beginning to understand many of the technical aspects of randomized controlled trials in education and the balances and trade-offs that must be attended to. The authors in this book have contributed generously and insightfully in the spirit of increasing our awareness and understanding of the strategies and challenges of education science.

The first chapter, by Lynn Okagaki, Elizabeth Albro and Jacquelyn Buckley, describes the genesis and development of the Institute of Education Sciences, the flagship in the development of scientific understanding of education in the United States. Their chapter traces the idea of scientific approaches to education over the last forty years, describes the Institute’s philosophy, and provides examples of
recent projects. The chapter provides an excellent foundation for understanding many of the issues addressed throughout the rest of the book.

Donald Christie and Andrew Pollard describe the development of the Teaching and Learning Research Project in the UK, one of the most ambitious attempts ever undertaken to develop research capacity and deep evidence across an entire field. They look at the history and development of the initiative, but also spend time considering the epistemological implications of the endeavor. The difference in approach to the issues of education science represented in these two chapters is striking, highlighting very nicely how context affects how developmental efforts play out.

The third chapter, by Ian Menter, looks at the way teacher education is affected by the notions of education science. Menter, one of the leading thinkers in teacher education in the UK, calls for a broad approach to educational research, including the active involvement of teachers, researchers, and policymakers. While welcoming the development of education science he also recognizes that the way it is put into practice (in both senses of the term) is critical for the improvement of educational processes.

The next three chapters engage deeply and directly with education science methodology, providing insights into the “stuff” of education science. Bob Slavin tackles the matter of producing systematic reviews in the fourth chapter. He looks at a number of the key problems and develops clear and incredibly useful suggestions for how the impact of these problems can be minimized. In the following chapter Carole Torgerson and David Torgerson look at why randomized controlled trials are considered to be so helpful, including a brief historical review of their development in education. Many of those who consider themselves opposed to randomized trials will find a great deal of food for thought in this chapter, which lays out the issues in a clear and non-dogmatic way. The final methodological chapter is a contribution by Ben Styles, who reviews randomized controlled trials—and the statistical logic underpinning them—in some detail. Again, this chapter demonstrates the sophisticated thought that supports experimental design. It is also noteworthy that these writers, some of the world’s leading proponents of education science, take care to point out that randomized controlled trials are a powerful method, but cannot be applied universally or uncritically.

The final four chapters move into a discussion of the limits of education science. Martyn Hammersley investigates the notion of evidence, asking what forms it can take and how it can be brought to bear on educational practices. He is particularly interested in judgment and trial-and-error as essential and valuable components of informed professional practice. In chapter 8 Patti Lather looks at the whole idea of science and scientism, suggesting that this epistemological position can usefully be considered as having old and new variants. She explores the contours of each of these in some detail, and presents a challenging conclusion that there is much to be learned from science and scientism, but that this requires recasting the terms and creating a non-reductionist scientificity.

Alison Phipps, in chapter 9, discusses the role of Humanities based investigation into education, and the dangers of falling into a kind of meager knowing if any one
form of educational research becomes dominant. She plays with ideas of sensual forms of knowledge and the place of imagination, arguing that science can contribute, but not without conscience. In the final chapter, I take on the task of situating the debate, suggesting that we cannot focus on education science as a purely methodological and epistemological debate while disregarding the changing political economy of educational research. Not only the practices of research are at stake, I suggest, but the frames for those practices.

My hope is that any reader, coming into this book from any perspective on the key issues of education science, will find something to stimulate them, to challenge them, and perhaps even inspire them to get involved in shaping the processes that are so profoundly transforming our field.

REFERENCES


LYNN OKAGAKI, ELIZABETH R. ALBRO
AND JACQUELYN A. BUCKLEY

1. INSTITUTE OF EDUCATION SCIENCES –PUTTING SCIENCE BACK INTO EDUCATION RESEARCH

In 1971, at the request of the President’s Commission on School Finance, the RAND Corporation began a review of education research in the United States. At that time, policymakers were looking for answers to basic policy questions – what programs and policies should be implemented to improve education outcomes for U.S. students. As part of their review, the authors considered a wide range of research including case studies of schools and school systems, correlational studies of the relations between variation in school resources and education outcomes, small-scale experiments comparing different instructional approaches, and quasi-experimental evaluations of programs. After reviewing the quality of available studies, they concluded that education researchers had produced very few answers to the practical questions of education decision-makers and practitioners. In the conclusion of their 222-page report, the authors wrote:

[Education research has tended to be small in scale, narrow in scope, diffuse, maldistributed, and lacking in focus. By comparison with other major sectors, the amount of research activity devoted to educational problems is surprisingly small. For example, the amount of resources allocated to agricultural research and development is more than four times as large, and health research is allocated more than 13 times as much . . . The body of educational research now available leaves much to be desired, at least by comparison with the level of understanding that has been achieved in numerous other fields. (Averch et al., 1972, p. 157)]

Fast-forward to 1999, when the National Research Council published a report on the state of education research in the United States. The panel concluded:

One striking fact is that the complex world of education—unlike defense, health care, or industrial production—does not rest on a strong research base. In no other field are personal experience and ideology so frequently relied on to make policy choices, and in no other field is the research base so inadequate and little used. Comparatively little research is funded, and the task of importing even the strongest research findings into over a million classrooms is daunting. (National Research Council, 1999, p. 1)

After a quarter of a century, the conclusions drawn about education research in the United States were remarkably unchanged. Why did we not make more progress in producing research to improve education and enhance student outcomes?
In this chapter, we present our view of the historical context and rationale leading up to the creation of the Institute of Education Sciences (Institute) and its approach to education research. The Institute is the primary research, evaluation, and statistics office within the U.S. Department of Education and was established in 2002. Its mission is decidedly pragmatic – to provide trustworthy information that education policymakers, leaders, and practitioners can use to improve the outcomes of education for all students, particularly those at risk of failure – and its approach to education research should be understood in relation to this mission. We begin by describing the Institute’s perspective on education, suggest that the tools for advancing education (e.g., experimental designs) have been available, yet rarely used, for several decades by education researchers, and observe that similar tools for research have been mainstays of social scientists in several disciplines. We then explain the legislative mission of the Institute and delineate what we mean by evidence-based education. Finally, we describe the Institute’s approach to education research.

DÉJÀ VU

Within the field of education research, we have recently seen much attention devoted to discussion of methodological choices and the implications of those choices (e.g., Angrist, 2004; Cook, 2002; Mosteller & Boruch, 2002; National Research Council, 2002, 2004). In particular, there has been much discussion of what kinds of methods enable us to answer questions of cause and effect. Worrying about these questions regarding establishing causality only makes sense in the context of one’s conception of the purpose of education; this is our starting point.

EDUCATION AS INTERVENTION

The Institute’s perspective on the purpose of education, in a nutshell, is that education is an intervention. Education is intended to provide students with the knowledge and skills to live meaningful lives and to contribute to society. A good education increases the choices that students will have in life and enables them to take better advantage of the opportunities that come their way. There have been a number of occasions in which attention in the United States has focused on the effectiveness of the education system for accomplishing national goals. Let’s consider two modern examples. First, after World War II, the private sector began expressing concern for a better-trained workforce in science and technology (Frechtling, Sharp, Carey & Vaden-Kiernan, 1995). Companies such as the General Electric Company and Westinghouse initiated programs in conjunction with universities to improve the science and technology backgrounds of high school science teachers. Soon after it was established in 1950, the National Science Foundation began sponsoring the first of many professional development summer institutes for high school science teachers. The great catalyst, however, that ignited attention on the quality of U.S. education was the Soviet Union’s space program. After the Soviet Union announced that Sputnik had orbited the earth, national attention turned to science education. Where the Soviet Union’s education system
was deemed to have successfully produced the scientists and engineers responsible for its accomplishment, the U.S. education system was found lacking (Douglass, 1999). Sputnik evoked a call for greater federal involvement in education that led to the National Defense Education Act of 1958. The primary purpose of the Act was to improve student learning in the sciences, mathematics, and foreign languages. The education system was the instrument that would produce scientists, engineers, and mathematicians that the country needed to meet the challenge to engage in the exploration of space.

Education again came into the spotlight when President Lyndon Johnson announced the War on Poverty to address the needs of some 35 million Americans living in conditions of poverty at that time. In this case, education was deemed a key tool for fighting poverty. President Johnson envisioned what he called the “Great Society.” In a speech at the University of Michigan, President Johnson observed,

A third place to build the Great Society is in the classrooms of America. There your children’s lives will be shaped. Our society will not be great until every young mind is set free to scan the farthest reaches of thought and imagination. We are still far from that goal . . . We must seek an educational system which grows in excellence as it grows in size. This means better training for our teachers. It means preparing young to enjoy their hours of leisure as well as their hours of labor. It means exploring new techniques of teaching, to find new ways to stimulate the love learning and the capacity for creation. (Johnson, 1964, p. 3)

In his treatise on the history of education in the United States Cubberley (1919) wrote:

With the passage of time, the growth of our Nation, and the extension of the suffrage to more and more diverse elements in our population, we have come to see clearly that an uneducated citizenship is a public peril, and to insist more strongly than before on the exercise of this fundamental right of the State. (Cubberley, 1919, p. 488).

He went on to argue:

In consequence it may now be regarded as a settled conviction of our American people that the provision of a liberal system of free non-sectarian public schools, in which equal opportunity is provided for all...is not only an inescapable obligation of our States to their future citizens, but also that nothing which the State does for its people contributes so much to the moral uplift, to a higher civic virtue, and to increased economic returns to the State as does a generous system of free public schools. (Cubberley, 1919, p. 495)

Education, as conceptualized by the nation’s leaders and policymakers, is a purposeful enterprise intended to enable the country to enhance productivity, ensure democracy, maintain national security, and promote the general welfare of its citizens.
RESEARCH TO IMPROVE EDUCATION

Working from the premise that education is an intervention that should enhance a variety of outcomes for students, we come to the question of how can research help education leaders and practitioners improve the quality of education? In an essay on American education research, economist Joshua Angrist (2004, p.198) observed, “The bread and butter issues in education research are questions about cause and effect in schools.” That is, questions about what works to improve education outcomes. At the heart of recent discussions about education research methodologies are questions about the types of research methods that are appropriate for answering questions of causality. More specifically, the division is between those for and against the use of randomized controlled studies to answer questions about the effects of education interventions (e.g., curricula, instructional approaches, policies). Given the constraints of this chapter, we are not reviewing arguments for the use of randomized trials for addressing questions about the effectiveness of education programs, practices, and policies; others have eloquently set forth such arguments (e.g., Angrist, 2004; Boruch, De Moya, & Snyder, 2002; Cook, 2002; Gueron, 2002). However, we note here that major reports from the National Research Council (2002, 2004) have concluded that randomized field trials are the most appropriate designs to utilize for research questions that seek answers to what works in education. According to the American Educational Research Association,

When correctly implemented, the randomized controlled experiment is the most powerful design for detecting treatment effects. The random assignment of participants to treatment conditions assures 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 to the pretreatment characteristics. (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007, p. 11)

As discussed below, the Institute values and supports research that addresses questions other than the what-works question, including exploratory research and development work. That said, the Institute does take the position that research to identify what works, for whom, and under what conditions is critical for improving the quality of education in the United States. Given the above recommendations by the National Research Council and American Educational Research Association, it is important to understand the history of experimental designs in education research and other social sciences.

FALL OF EXPERIMENTAL DESIGNS IN EDUCATION RESEARCH

Researchers in agricultural sciences began using randomized controlled field trials in the United States in the 1920s and 1930s (Nave, Meich, & Mosteller, 1999). Industry and medicine soon followed. The methodology was eventually picked up by education researchers. For example, in the 1960s and 1970s a series of
evaluations of Head Start and Head Start-related demonstration programs were implemented. One of the most widely cited of these evaluations is the High Scope/Perry Preschool Project. In this evaluation, children were randomly assigned to the treatment or control condition. The treatment included a half-day, center-based preschool program for the children and home visits. The initial evaluation and its subsequent follow-ups have shown that children in the High Scope group were more likely to graduate from high school; by age 40, they were more likely to be employed, to own a car, and own their own home. Although the Perry Preschool Project was a small evaluation focusing on a sample of only 123 African American preschool children (Schweinhart, 2004), it continues to be cited as evidence in support of the effectiveness of early childhood intervention programs.

By 1980, however, when Boruch and Cordray published a 419-page report in response to legislation requiring the Commissioner of Education to review the state of evaluation practices for education programs, incorporating experimental designs into the evaluation of education was still relatively rare. Boruch and Cordray examined over 400 reports on evaluations of local, state, and federal education programs and conducted interviews with education policymakers, leaders, and staff at the local, state, and federal levels. Among their conclusions, they wrote:

High quality evaluation designs for estimating the effects of programs on children are not used often. We recommend that the Department [of Education] authorize explicitly the use of randomized field experiments to plan and evaluate new programs, new variations on existing programs, and program components. (Boruch & Cordray, 1980, p. 4)

Given the National Research Council’s similar conclusion about the state of the field in 1999, one wonders why experimental designs were not more embraced by education researchers in the latter decades of the 20th Century. A number of scholars have considered this question. Cook (2001) observed that in the 1970s education researchers began to describe schools as “complex social organizations” and to adopt qualitative methodologies for examining and evaluating the management of education organizations and the implementation of education programs. Many leaders in education evaluation explicitly rejected experimental methods for determining causality in education settings (e.g. Cronbach, 1982; Eisner, 1975; Fetterman, 1982). Eisner wrote:

Those familiar with the evolution of the evaluation field already know that it has been significantly influenced by the assumptions and procedures employed in doing educational research. And education research in turn took as its model the natural sciences and had as its aspiration the development of theory and methods that would make educational practice scientific . . . Yet, scientific procedures are not the only forms through which human understanding is secured and scientific methods are not the only ways through which human influence can be confidently created. What I shall do in this paper is to suggest, perhaps not so much an alternative, but surely a needed supplement to the use of scientific procedures for describing,
interpreting and evaluating educational settings. I call this new, non-scientific approach to educational evaluation an approach that requires educational connoisseurship and educational criticism. (Eisner, 1975, p. 5)

Since that time, relatively little work has included experimental designs. Nave, Meich, and Mosteller (1999) reported that less than one percent of education dissertations or studies in Education Resources Information Center (ERIC) Abstracts were randomized experiments. Hseih and colleagues (2005) examined articles on intervention studies from four educational psychology journals and from one general education research journal that were published in 1983 and from 1995 through 2004. In 1983, 47 percent of the educational psychology articles and 32 percent of the articles in the American Educational Research Journal (AERJ) involved random assignment to conditions. By 2004, the percent of randomized experiments had dropped to 26 percent among education psychology articles and 4 percent of AERJ articles.

In a short essay on trends in education evaluation from the 1960s to 1990, Ernest House wrote:

Between 1965 and 1990, the methodology, philosophy, and politics of evaluation changed substantially . . . Evaluation moved from monolithic to pluralist conceptions, to multiple methods, measures, criteria, perspectives, audiences, and even multiple interests. Methodologically, evaluation moved from a primary emphasis on quantitative methods, in which the standardized achievement test employed in a randomized experimental control group design was most highly regarded, to a more permissive attitude in which qualitative research methods became acceptable. Mixed data collection methods are now routinely advocated in a spirit of methodological ecumenicalism. (House, 1990, p. 24–25)

According to House, the move away from randomized experimental designs was the result of multiple factors. The 1960s and 1970s were decades in which the United States was experiencing a cultural upheaval. This was a time in which the metaphor of the United States as a melting pot in which people from diverse backgrounds blended into one was branded as being antiquated and in current terms, politically incorrect. We became a country in which the voices, the values, and the perspectives of diverse groups within the country were heard as distinct melodies. People marched for the Civil Rights movement, against the Vietnam War, and for women’s liberation. The acknowledgement that people from different groups have different perspectives and different goals had implications for program evaluation. Using qualitative methods to capture the richness, nuances, and diversity of views was a reasonable choice.

Concurrent with these cultural events came the recognition of the importance of capturing richer and more sensitive data – of the need for including multiple measures, for obtaining process data – and for developing more detailed conceptualizations of intervention theories of action. Ultimately, education researchers developed more sensitive approaches to collecting and analyzing descriptive data – both from qualitative methodologies and the development of sophisticated data modeling.
techniques, such as LISREL. The methodological tools available to education researchers became much more diverse. But experimental methods were generally still missing from the education research toolbox.

EXPERIMENTAL RESEARCH IN THE SOCIAL SCIENCES

The use of random assignment and experimental designs to draw conclusions about causal effects is common in other social sciences. Cook (2002) observed that social scientists across a number of disciplines have relied on experimental designs to answer causal questions.

The superiority of random assignment for drawing inferences about the consequences of planned change attempts is routinely acknowledged in … micro-economics, psychology, criminology, prevention research, early childhood education, marketing and those part of political science and sociology concerned with improving opinion surveys. It is also acknowledged in all of the elementary education method textbooks we have consulted. However, random assignment is relatively rare in educational research, especially for assessing the impact of educational interventions of obvious policy relevance. (Cook, 2002, p. 177)

Psychology has relied upon experimental methods over the past century to examine how to improve teaching and learning. The empirical tradition in the study of the psychology of learning has roots both in the experimental tradition of the laboratory, as well as in the naturalistic observational tradition. The experimental tradition emerges out of early German psychology and the laboratories of researchers like Wundt, and can trace its ancestry through American behaviorism to modern cognitive science. In his historical account of the emergence of the discipline of cognitive science, Gardner (1985) described how the empirical tradition of experiment in psychology played a key role in the construction of cognitive science. As the discipline became increasingly dissatisfied with what could be learned about learning by studying observable behavioral outcomes only, cognitive psychologists designed experiments intended to elucidate the mental representations and processing of information that underlay the behavior. Over the past forty years, cognitive scientists have made fundamental discoveries about critical aspects of learning, and have used knowledge emerging from psychological experiments to inform education practice. For example, more than a hundred experimental studies have demonstrated that spacing practice over time improves learning as compared to massing practice (i.e. cramming for a test does not improve long-term memory) (Cepeda et al., 2006). Recognizing that in many of these early experimental studies, learners were asked to remember isolated units of information and for relatively short periods of time, current researchers are using experimental designs to ask whether the same spaced exposure to knowledge is equally effective when students are being asked to master more complex sets of information and for longer periods of time.

Whereas much psychological research has not been conducted in schools, the use of experimental design in school-based studies is common in fields such as
prevention science and economics. In a recent special issue of the *American Journal of Preventive Medicine*, Henry and colleagues (2004) described a multisite violence prevention project in which middle schools were randomly assigned to one of four conditions in a full 2 x 2 design. In their explanation of the design of this study, the authors discussed “the necessity of a true experimental design to assess efficacy . . . such designs are critical for producing the scientifically valid knowledge needed to advance the field of violence prevention” (Henry et al., 2004, p. 13). Throughout this issue, researchers noted both the challenges and successes of carrying out a cluster-randomized trial, but nowhere do the authors conclude that experimental research is inappropriate for answering questions of efficacy. In fact, in its statement on standards of evidence for determining whether prevention programs are effective, the Society for Prevention Research indicated that “[f]or most kinds of interventions, random assignment (of sufficient sample size without significant pretest differences) is essential” (Society for Prevention Research, 2004, p. 3).

Economists have often used experimental designs in their work, including evaluations of the impacts of school-based interventions (e.g. Angrist, Lang, & Oreopoulos, 2006; Decker, et al., 2004; James-Burdumy, Dynarski, & Deke, 2008; Lavy, 2002). For example, Dynarski and his colleagues (1998) evaluated the effects of sixteen dropout prevention programs on 6,000 middle and high school students randomly assigned to either a treatment or control group. Students in the treatment group participated in one of 16 programs. For the middle school students, participating in a high intensity (e.g. full-day) prevention program improved the likelihood of grade promotion and had lower dropout rates. Receiving a General Educational Development certificate is generally recognized as being equivalent to receiving a high school diploma in the United States, and high school students participating in the dropout prevention programs were more likely to receive these certificates as compared to students in the control condition. On the other hand, the experimental findings demonstrated that students attending alternative high schools were no less likely to drop out than students attending typical high schools. As another example, economist Paul Decker and his colleagues (2004) conducted a randomized controlled trial to determine if Teach for America teachers improve – or, at minimum, do not hinder – student achievement compared to business-as-usual (i.e., compared to whatever teacher students would have been assigned to if their school were not participating in a study of the Teach for America program). Teach for America is a program that recruits recent college graduates and working professionals with strong academic records and provides them with intensive summer training and ongoing professional development in exchange for a two-year commitment to teach in high-need (e.g. high poverty) rural or urban schools. In this evaluation of 100 teachers and about 2,000 students in 17 elementary schools across six regions of the country, students who had Teach for America teachers learned more in math relative to students in other classes, and there was no statistically detectable difference in reading scores.
Experimental design has been a prominent research tool for determining causal effects in the social sciences, but has remained relatively unused in education research. Perhaps, it is also fair to say that asking questions about the effectiveness of programs for improving education outcomes has not been central to education research in the last few decades. Using high quality experimental designs to identify those interventions that substantially improve education outcomes can increase the quality of education and improve outcomes just as it has in other fields.

EDUCATION IN THE UNITED STATES

In the United States, most will agree that our education system needs to be improved and that we can do a better job of educating our students. For instance, U.S. students continue to lag behind their international peers in mathematics and science. According to the results of the Program for International Student Assessment 2006, which reports on the scientific literacy of 15-year-olds in 57 education jurisdictions, including 30 member countries of the Organization for Economic Cooperation and Development (OECD), the average U.S. scientific literacy score was below the overall average of the 30 OECD countries (Baldi, Jin, Skemer, Green & Herget, 2007).

On a national level, our outcomes are just as troublesome. Our dropout rate has hovered around 20 percent since 1970 (Heckman & LaFontaine, 2007). In addition, educational inequity persists along racial, ethnic, and socioeconomic lines. For example, since 1970, minority children have consistently scored 30-40 points lower than their peers on the NAEP reading assessment (Labov, 2003; NAEP website http://nces.ed.gov/nationsreportcard). Furthermore, around 30-40% of minority youth are not completing high school and that has remained relatively stable over time (Heckman & LaFontaine, 2007).

These outcomes are stunning considering that the federal government spends nearly 120 billion per year on education programs, and federal education expenditures represent only a small fraction of the total education expenditure. Despite the dedicated fiscal support, we clearly are not achieving educational excellence. Education leaders and practitioners work hard to help their students; they need answers from education researchers as to what works and for which students. The U.S. Congress recognized that improving student outcomes would require a fundamental shift in education research in our country. This was the impetus behind the creation of the Institute of Education Sciences.

INSTITUTE OF EDUCATION SCIENCES: LEGISLATIVE MISSION AND CONTEXT

The Institute was established through the Education Sciences Reform Act of 2002. The mission of the Institute is to provide reliable information on the (a) condition and progress of education in the United States, (b) education practices that improve academic achievement and increase access to education opportunities for all students, and (c) the effectiveness of Federal and other education programs. By law the Institute has a decidedly applied focus. Shortly after its creation, the Institute asked education decision-makers—superintendents and other local education officials, chief state school officers, state higher education officers, governors’
education policy advisors, state legislators, congressional staff members, and executive directors of national education associations—what questions they wanted education researchers to answer. The highest priority for education decision-makers was research that would help them identify programs, practices, and policies that would improve student achievement, particularly in reading, mathematics, and science (Huang et al., 2003). Those who are responsible for educating students want to know which programs and practices they can implement in their schools in order to improve education outcomes for their students. They were critical of what they perceived to be the impractical focus of existing education research. According to one superintendent:

There may be less than one percent of the existing research that’s really meaningful to teachers. Much is for researchers, for getting funding, for career advancement, or for advocacy. . . . I don’t want theories. Teachers need strategies, practices. Give them things that can help teaching and learning, things that can help kids. (Huang, et al., 2003, p. 16).

The Institute was given the job of changing federally-funded education research in the United States so that it would provide trustworthy information to help local, state, and federal education policymakers, administrators, and practitioners improve the quality of education and thereby improve valued student outcomes.

**EVIDENCE-BASED EDUCATION**

In a speech at Northwestern University, the Institute’s first director Grover J. Whitehurst articulated the Institute’s position on grounding education decision-making in scientific evidence.

Our guiding premise is that scientific research and evaluation, linked with systematically collected and utilized data on education performance, is the key to progress in education. Indeed, we assume that evidence-based practice and policy is the best—and perhaps the only—way to produce continuous improvement in education outcomes. (Whitehurst, 2004, p. 2)

Figure 1 is a schematic depiction of the Institute’s conceptualization of the evidence-based decision-making process. Scientific research, evaluation, and statistics along with performance data constitute the empirical evidence used to inform decisions. Performance data refers to the data that education systems collect to monitor the degree to which the programs they deliver are achieving the intended outcomes (e.g., percentage of students reading at grade level, on-time graduation rates). The Institute’s definition of scientific research, evaluation, and statistics accords with the National Research Council’s view. In a 2002 National Research Council report, the panel concluded that all scientific research, including research in the education sciences, must “pose significant questions that can be investigated empirically . . . link research to relevant theory . . . use methods that permit direct investigation of the questions . . . provide a coherent and explicit chain of reasoning . . . yield findings that replicate and generalize across studies . . .
and disclose research data and methods to enable and encourage professional scrutiny and critique” (National Research Council, 2002, pp. 3–5). By contexts and constraints, we mean those local conditions that must be considered when making decisions (e.g., types of students and teachers, resources, local infrastructure for delivering programs). For example, upon learning that an evaluation of a dropout prevention program showed that high schools implementing the new program had better attendance rates, reduced dropout rates, and better graduation rates, a superintendent might first consider the degree to which the students, teachers, and schools in the evaluation are similar to those in her district and judge whether the findings of the evaluation are likely to generalize to her district. The superintendent is also likely to consider the cost of the program and whether or not the program is feasible for implementing in the district before making a decision to buy the new program. These are all local conditions that are important considerations in evidence-based decision making.

**Figure 1.1. Evidence-based decision-making process**

**BREADTH OF EDUCATION RESEARCH AT THE INSTITUTE**

The mandate to supporting research that will generate empirical evidence to improve the quality of education in the United States is broad, and the work of the Institute encompasses a wide range of research activities, including descriptive studies, development and innovation, evaluation of the effectiveness of education programs and policies, and assessment research. All are needed to improve the quality of education. In what follows, we provide examples of the types of work
that are supported by the Institute, the variety of research methods that are employed, and how the work contributes to transforming education into an evidence-based field.

Descriptive studies. The National Center for Education Statistics (NCES), the oldest and perhaps most well known of the Institute’s four centers, conducts descriptive studies that enable it to report on the condition and progress of education in the United States. Among its many responsibilities, NCES fields the National Assessment of Education Progress (NAEP), which provides a national picture of student achievement in reading, mathematics, science, and other major curricular domains. Utilizing national surveys and administrative data, NCES reports on a wide range of issues relevant to the characteristics of schools, students, faculty, education programs, school finance, and other indicators of the quality and extent of education opportunities offered in the country. The methodological approaches that are utilized to describe the condition and progress of education are varied and include surveys, large-scale assessments, analyses of administrative records, structured interviews, classroom observations, and case studies. Descriptive research of this nature is critical to the transformation of education into an evidence-based field. For example, tracking student achievement through national assessments and state-level achievement tests, as well as monitoring other valued education outcomes (e.g. high school graduation rates, transition to postsecondary education) helps education leaders determine if the education system is achieving what the country expects it to accomplish. Analyses of these data can highlight successes and problems and can contribute to the generation of hypotheses relevant to the identification of possible solutions to problems.

Another type of descriptive research is the examination of correlations between factors, such as education programs and student outcomes. Through the Institute’s education and special education research centers, the Institute provides grants to external researchers (e.g., researchers at academic institutions) to conduct research. The Institute funds five basic types of research:

– exploratory studies;
– projects to develop new and innovative interventions;
– evaluations of the efficacy of interventions under limited conditions;
– evaluations of the impact of interventions when they are implemented at scale;
– measurement development and validation.

Under the exploration category, the Institute supports projects that are intended to explore the relations between malleable factors (i.e., factors that can be changed and are potential targets of interventions, such as child characteristics or education practices) and education outcomes; or examine factors and conditions that may mediate or moderate the relations between malleable factors and education outcomes.

Within this category, a number of different research approaches are taken. Investigators may conduct secondary analyses of administrative data to examine the associations between different education programs and student outcomes. Using detailed observational data, researchers may link specific education practices to student outcomes to identify malleable factors that are potentially amenable to
intervention. When sufficient numbers of studies are available to support a probing meta-analysis, researchers may use meta-analytic techniques to identify issues that potentially affect or moderate the results, such as quality of the research design and type of implementation. The common denominator in exploration studies is that they are intended to generate hypotheses about education interventions – pinpointing malleable factors predictive of education outcomes and potentially amenable to intervention, identifying promising interventions that should be rigorously evaluated, or finding factors that may mediate or moderate the effects of education interventions. This type of research is sometimes called “translational” research because its ultimate purpose is to be translated into interventions.

As examples of the importance of exploratory research, consider the following study carried out by researchers at the Center for Analysis of Longitudinal Data in Education Research (CALDER), a national research and development center supported by the Institute. Using administrative state datasets in which student and teacher data are linked, a team of CALDER researchers compared the quality of the teachers and principals in high poverty schools in North Carolina to the quality of teachers and principals in schools serving more advantaged students (Clotfelder, Ladd, Vigdor, & Wheeler, 2007). The findings were clear and mirrored findings reported by others (e.g. Betts, Reuben, & Danenberg, 2000; Lankford, Loeb, & Wyckoff, 2002). When compared to teachers in low-poverty schools, teachers in high-poverty schools have less teaching experience, attended less competitive undergraduate institutions, have lower licensure test scores, and are less likely to be board certified. A similar set of quality metrics for principals was gathered and analyzed, and the same pattern of results was identified. Clotfelder and his colleagues concluded that the least qualified teachers and principals typically instruct students with the most need. These findings identify a problem – the maldistribution of high quality teachers – and raise the question of whether policies can be implemented to redistribute the most highly qualified teachers into schools with the most need; the results of this study can be translated into new policies that can be tested in subsequent research.

What policies might be used to redistribute high quality teachers? Recently, New York City attempted to address the typical maldistribution of instructional resources by changing teacher qualification requirements. By abolishing temporary licenses, creating alternative certification routes, and developing a Teaching Fellows program, the city has been able to radically alter the characteristics of teachers within New York City. Another group of CALDER researchers (Boyd et al., 2007) reported that the gap between the qualifications of New York City teachers in high- and low-poverty schools narrowed substantially between 2000 and 2005. This reduction in the qualification gap was associated with improvements in student achievement scores.

Through the Institute’s Cognition and Student Learning program, cognitive scientists have conducted primary descriptive work to identify processes that may be targets for new interventions. For example, Robert Siegler and his colleagues (Laski & Siegler, 2007; Opfer & Siegler, 2007; Siegler & Opfer, 2003) conducted a series of studies examining whether variations in performance on number-line tasks predicted broader individual differences in estimation skills. At the same
time, Siegler also asked whether children who consistently rely on accurate, linear, representations of numerical magnitude have an easier time learning arithmetic skills than children who sometimes or always rely on a logarithmic representation. Through these studies, they identified numerical estimation as a foundational skill for mastering basic arithmetic (Booth & Siegler, 2008; Seigler & Booth, 2004; Seigler & Booth, 2005).

Everyone they tested – preschool children, upper elementary school children, and college students – estimated the value of numbers using a logarithmic scale, as opposed to a linear scale; the only difference was a matter of magnitude. Young children struggled with the 1-10 number line; adults with a number line that reached into the thousands. Examining existing empirical literature, Booth and Siegler (2006; 2008) ascertained that individuals with a linear understanding of the relationship of numbers to a number line performed better at arithmetical functions. Further, they found that children from impoverished backgrounds tended to perform less well on both the estimation and arithmetic function. Siegler has subsequently used this descriptive research as the basis for developing interventions for preschool children and elementary-school children that are intended to help them master the foundational understanding that when numbers increase by one along a number line, the unit of “one” is the same regardless of whether the starting number is 2 or 250. As one example, Siegler and his colleagues (Siegler & Ramani, 2008) designed a simple board game for four-year-old children play and in their initial evaluations have found that it enhances basic mathematical understanding for at-risk young children.

Development and innovation. Development of new and innovative education interventions is critical to improving education. Since 2004, when the Institute established the five types of research projects that it funds through its regular research grant programs, over half of the funded grants have been projects to develop new interventions, including curricula, instructional approaches, teacher professional development programs, education technology, and systems-level interventions (National Board for Education Sciences, 2008). Developing new interventions that have the potential to improve education outcomes requires a solid knowledge of existing education practice to recognize how they could be improved and to understand the constraints of the system to know what is feasible for implementation in a new intervention. It also requires the creativity to conceptualize innovative approaches that are sufficiently different from existing practices potentially to improve outcomes over and above what current practices produce.

In typical development projects, researchers incorporate multiple methodological approaches, including procedures for designing and creating the prototype, for determining if the intervention functions as intended (e.g. Does the software program crash when students use it? Can students use the technology independently and do they understand the information in the way that it is presented?), and for determining if the intervention can be implemented within the constraints of an actual education program. Within a typical project, researchers might include
detailed observations of individuals, interviews, focus groups, think-aloud protocols, as well as small experiments and quasi-experiments.

With Institute support, Scott Baker and his colleagues developed an intervention designed to support the literacy skills of first graders (Santoro, Chard, Howard, & Baker, 2008). Building upon empirical work on the development of reading skills and substantial knowledge of reading instruction, Baker planned a 15-unit/15 week story read-aloud intervention, which was designed to improve beginning readers’ vocabulary knowledge and comprehension skills. In creating the intervention, the researchers focused on activities intended to support young readers’ understanding of both narrative and expository text and to enable teachers to use the story read-aloud framework in a much richer way than found in typical instruction. Development of the initial prototype took several months. The research team incorporated input from teachers as they developed a framework that would address state standards for language arts and reading comprehension, as well as state standards for science and social studies; incorporate expository and narrative texts; and move instruction from teacher-directed interactions to more independent student responses.

To develop the curriculum, researchers scoured children’s books to identify those that met their criteria, constructed lessons that addressed both vocabulary learning and comprehension, and prepared professional development modules. As the researchers developed and pilot tested components, feedback was obtained from experienced teachers and revisions were made based on this feedback. After all of the components had been created, the team field-tested the full 15-unit, 15-week intervention with several teachers and the students in their classrooms collecting student pre- and post-intervention data, observing in classrooms how the intervention unfolded when delivered by teachers, and conducting debriefing interviews with teachers who had delivered the intervention.

As quickly became apparent, the researchers needed to revise the curriculum because teachers could not complete all of the scripted activities within the allotted time for each lesson, the social studies units and books were too difficult for the first graders, and covering one unit a week was too fast a pace. Thus, the team decided to remove the social studies units from the curriculum, and redesigned the delivery of the 9 science units for implementation over 18 weeks. In addition, an examination of the student data indicated that although the embedded vocabulary instruction appeared to benefit students who started with a strong vocabulary, it was not intense enough to help lower-performing students. Consequently, the team decided to develop a series of vocabulary “booster” lessons for students who needed additional instruction. A second field test was conducted to assess the implementation of a shorter main intervention that was coupled with vocabulary booster lessons for targeted students. This version produced stronger results, and the team is currently engaged in an efficacy study to test whether participating in the Story Read-Aloud intervention with or without the booster affects student literacy outcomes.

The Institute also encourages innovation in education. For example, despite considerable research on technology use by children and adolescents and the development of technology for young people, many have argued that we have not
yet exploited the full potential of electronic media for educational purposes (e.g. Federation of American Scientists, 2006; Wideman et al., 2007). In particular, there have been relatively few attempts to use advanced, high-resolution multimedia gaming environments as contexts for the instruction of basic academic skills. In general, existing education technology products are not ones that students would naturally gravitate to outside of school; they typically lack high quality graphics and sounds, sophisticated user interface, a reward structure that cultivates a strong sense of motivation in users, and engaging activities that maintain the user’s attention. Consequently, the Institute asked for proposals to create innovative education technology that utilizes rich multimedia gaming environments as a context for learning academic content.

In 2008, two new research and development centers on instructional technology were established. One center is creating a science and literacy intervention that will be delivered on the Nintendo Dual-Screen portable gaming environment. The system includes multi-player communication capability and utilizes gaming techniques, such as a story framework, and competition among players, to increase the engagement and persistence of students with the academic content. The other center is developing a gaming environment that will foster mastery of pre-algebra and algebra concepts. Both of these centers are using current strengths of the technology-driven gaming environment to create innovative new products that hold the promise of engaging and supporting the achievement of underperforming middle and high school students.

Evaluation. In “Scientific Research in Education” the National Research Council commented on the progression of science:

An area of research that, for example, does not advance beyond the descriptive phase toward more precise scientific investigation of causal effects and mechanisms for a long period of time is clearly not contributing as much to knowledge as one that builds on prior work and moves toward more complete understanding of the causal structure. (National Research Council, 2002, p. 101)

Descriptive research on the state of education, exploratory work that can inform practice, and development and innovation are crucial components of the research enterprise that is needed to improve the quality of education. However, as the National Research Council observed, scientific research cannot stop with descriptive work. To move forward, we need to test causal effects and mechanisms.

The Institute has been most recognized (positively and negatively) for its support of randomized controlled experiments to evaluate the effect of interventions on education outcomes. For questions about “what works,” the Institute “has taken a clear and strong position that randomized trials, in which participants are assigned by lottery to experience two or more conditions, are the best method for determining what works” (Whitehurst, 2004, p. 9). In the context of a field that had turned away from experimental designs to answer causal questions, this position was a hard and abrupt change in direction. Through the Institute’s National Center
for Education Evaluation and Regional Assistance, the Institute conducts large evaluations of the impact of federal and other education programs. By 2008, 26 evaluations had been completed or were underway; 22 of these evaluations were randomized controlled experiments. By way of comparison, Robert Boruch and his colleagues (2001) examined 144 evaluation studies awarded between 1995 and 1997 by the U.S. Department of Education’s Office of Planning and Evaluation Service. Of these 144 studies, 51 were intended to examine the impact of federal programs, but only 5 used an experimental design to measure that impact. Of the 84 program evaluations and studies planned by the Department of Education for 2000, 16 addressed the effectiveness of federal programs, but only one utilized a randomized controlled design to assess the effect (Boruch, DeMoya, & Snyder, 2002).

Through its two research centers – the National Center for Education Research and the National Center for Special Education Research – the Institute also supports evaluations of education interventions: (a) efficacy trials to determine whether or not interventions are effective under limited conditions and with specific types of students and (b) scale-up evaluations to determine whether or not interventions are effective when they are implemented under conditions that would be typical if a school district were to implement them (i.e., without special support from the developer or the research team) and across a variety of conditions (e.g., different student populations, different types of schools). Whereas over half of the research grants awarded by these centers have been projects to develop new education interventions, approximately one-fourth of the grants funded through their regular research competitions have been efficacy or scale-up evaluations (National Board for Education Sciences, 2008). It is only for answering questions about causal effects that the Institute prefers randomized trials, and even so, the Institute indicates that other methods are also acceptable. In descriptions of its education research programs, the Institute states:

Applicants may propose a quasi-experiment rather than a randomized trial when randomization is not possible or when the external validity of the quasi-experiment provides valuable information that is not obtainable from a randomized counterpart. Acceptable quasi-experiments will substantially minimize selection bias or allow it to be modeled. Possible approaches include regression-discontinuity designs, instrumental variable designs, or matched comparison groups designs in which equivalence is demonstrated between the intervention and comparison groups at program entry on the variables that are to be measured as program outcomes (e.g. student achievement scores). (Institute of Education Sciences, 2008)

The overall design of the study (experimental, quasi-experimental) and its implementation are essential requirements to answering the basic impact question (i.e., does the intervention produce a net positive impact on the outcomes of interest in relation to the program or practice to which it is being compared). However, the Institute expects the efficacy and scale-up evaluations funded through the two research centers to answer questions beyond the causal effect of the intervention. Researchers examine the fidelity of the implementation of the intervention, as well
as mediators and moderators of the effect of the intervention. In the context of the impact evaluation, researchers are encouraged to incorporate observational, survey, and other descriptive methodologies to assist in the identification of factors that may explain the effect or lack of effect of the intervention.

As an example of the type of evaluations that are funded Vadasy and Sanders (2008a, 2008b, 2008c) have recently completed three efficacy studies of Quick Reads, a repeated reading instruction intended to improve the word-level reading skills and fluency of below average readers. In each of these studies, the researchers examined how the effects of Quick Reads varied as a function of the age of the student (second and third graders or fourth and fifth graders), the characteristics of the tutor (para-educator or certified teacher), and the fidelity with which the intervention was delivered. Detailed classroom literacy instruction observations were completed in which observers measured instructional time across a range of literacy activities (e.g. phonological awareness, word study/phonics, spelling, oral language development, fluency building, text reading, comprehension, writing/language arts), as well as behavior management, evaluative feedback, transition time, other instructional activities, and students’ engagement in tasks.

Using these observational data, the researchers were able to test, for example, whether the amount of classroom time allocated to oral reading was associated with student gains, over and above gains associated with participation in Quick Reads tutoring sessions. They found that there was a significant, though small, correlation between the amount of classroom oral reading time and gains in word reading accuracy and fluency. In the most recently completed efficacy trial, Vadasy and Sanders (2008c) examined the relation between fidelity of implementation of the intervention and student outcomes. To gather fidelity information, researcher-observers completed observation forms that captured the tutors’ adherence to scripted Quick Reads protocols and their instructional behaviors.

Analysis of these observational data indicated that fidelity was associated with student gains – students of tutors whose instruction was closest to the ideal tutor behavior showed greater word reading and fluency gains. In each of these efficacy trials, then, the researchers gathered data that not only indicate whether the intervention worked to improve student reading outcomes, but also to help explain patterns of achievement. The purpose of efficacy and scale-up evaluations is to determine whether the target intervention improves education outcomes relative to what the comparison group experiences and for which students (or teachers or schools) and under what conditions. As much as possible, we want to know not only if something works, but also why it works (or does not work), under what conditions it works, and for whom it works. To address these questions, multiple methodological approaches are incorporated into the studies.

Assessment development and validation. Finally, the Institute supports research on assessment instruments. These projects include the development and validation of new student assessments, validation of existing assessments, development and testing of new techniques for assessment or analysis of assessment data in the context of state accountability standards and systems, as well as development and
validation of assessments used to certify or assess education professionals. A variety of methodological approaches are appropriate for this work.

One of the Institute’s current measurement projects is the development and validation of the Diagnostic Algebra Assessment System. Building on research on the types of misconceptions that underlie poor performance in algebra, Russell and his colleagues developed a set of items that could be used to identify if and when students were making errors during problem-solving due to holding specific misconceptions. For example, many students come to algebra with a misconception as to the meaning of the equal sign. Students often interpret the equal sign as meaning “do something” or the sign they see before the answer to a problem. They do not understand that the equal sign means that the expression on the left side of the equal sign must be equivalent to the expression on the right side of an equation. Without an accurate understanding of the equal sign, much of algebra instruction will not make sense. Studies were conducted to ensure that the items and the full test reliably distinguished among students who understood the material, those who held misconceptions, and those who made mistakes not based on underlying misconceptions. The assessment system was designed to be delivered online and to generate a summary report for teachers that indicates how the class as a whole performed and which students are struggling with which misconceptions. Supporting the development of these sorts of “just-in-time” assessment instruments holds promise that instruction can be shaped to meet the needs of individual learners.

FINAL OBSERVATIONS

Through its research, statistics, and evaluation activities, the Institute of Education Sciences supports a wide range of scientific research in education that address the practical needs of education policymakers, leaders, and practitioners. The primary purpose of these activities is to enhance the quality of education in the United States and improve outcomes for students.

We began our chapter with the premise that education is an intervention, that education is intended to achieve particular goals for students. In the United States, many believe that our education system needs to be improved. Achievement gaps exist between students from middle- and upper-income families and students from low-income families and between European-American students and African-American, Hispanic-American, and Native American students. In the 2005 NAEP, the most recent assessment of Grade 12 students, only two percent of U.S. students attained advanced levels of mathematics or science achievement; 39 percent of Grade 12 students scored below the basic level. In the 2007 NAEP, 18 percent of Grade 4 students and 29 percent of Grade 8 students scored below the basic level in mathematics. At Grade 4 scoring below the basic level means that the student is likely to miss problems such as using a ruler to find the total length of three line segments. At Grade 12 scoring below the basic level means that the student is unlikely to be able to solve problems such as finding the perimeter of a figure. The education system in the United States can do better than this for our students. Teachers, principals, and other education leaders are working hard to provide
students with the best education that they can deliver. They need better tools – curricula, instructional approaches, programs, assessments, technology – to increase the strength and potency of education. The Institute of Education Sciences was established to provide educators and education policymakers with the information and tools to improve education in our nation.

The Institute supports a broad range of research, statistics, and evaluation activities – including descriptive, exploratory, development and innovation, evaluation, and assessment projects. What methodologies are appropriate for this work? The Institute’s position is that the research methods used must be appropriate for the question asked. Different questions require different methodological tools. Across its research, statistics, and evaluation projects, the Institute supports projects that incorporate the use of multiple research, evaluation, and statistics methodologies.

Is the work that the Institute supports the only scholarship that is relevant to education? No, it is not. Historians and philosophers, for example, help us think about what education has been, what it could be, and perhaps, what it should be. But for the purpose of improving education today – of making education more potent and powerful for enhancing student outcomes – we need scientific research. We need research that enables us to understand the condition of education and to monitor the progress of our students. We need research that generates new hypotheses about education processes, about what can be changed or modified to improve outcomes, about why something might help some students but not others, and about policies that might improve the effectiveness of the education workforce. We need research to develop new and innovative approaches to education that has the potential to substantially increase learning and achievement. We need research that rigorously evaluates the impact of education interventions – programs, practices, and policies – to determine which ones are most effective for which students, under what conditions, and why. The Institute was created to do just that – to support a scientific enterprise that is dedicated to providing the information, the tools (e.g., assessments), and interventions (e.g., curricula, systems-level programs, policies) that will enable education leaders and practitioners to improve the quality of education in the United States and thereby enhance student outcomes.

NOTES
1 The opinions expressed are those of the authors and do not represent the views of the U.S. Department of Education.
2 The percentage of articles in AERJ that were randomized experiments in 1983 was not included in the article cited. This information was provided by Daniel Robinson on December 2, 2008.

REFERENCES


2. TAKING STOCK OF EDUCATIONAL RESEARCH 
AND THE IMPACT OF THE UK TEACHING AND 
LEARNING RESEARCH PROGRAM

INTRODUCTION – WHAT IS THE TEACHING AND LEARNING RESEARCH PROGRAM?

The Teaching and Learning Research Program (TLRP) has been by far the largest 
ever investment in educational research in the United Kingdom and certainly 
among the largest in the world. Its scale eclipses any other program of research 
ever before funded or managed by the UK Economic and Social Research Council 
(ESRC). TLRP has provided coordination and support for over 700 educational 
researchers in some 70 project teams and almost 20 initiatives of cross program 
thematic analysis across the four countries of the UK. The first projects began 
empirical work in 2000 and the last project is presently expected to end in 2012. 
The total budget by the end of 2008 was some £43 million having drawn together 
contributions from a wide range of UK government bodies, principally the Higher 
Education Funding Council for England to which credit is due for imaginative 
initiation of the scheme and the largest share of the funding. The scope and 
ambitions of TLRP have grown very substantially from its inception in 1999, when 
around £10 million was invested. It has attempted to support cutting edge research 
aimed at making a significant impact on the quality of teaching and learning across 
the life course and in a wide range of learning contexts in all parts of the UK. It has 
aimed to encourage the engagement of a wide range of social science perspectives 
in addressing research questions arising from the field of education and at the same 
time enhance research capacity. A major feature of the approach advocated by 
TLRP has been a fundamental commitment to user engagement, particularly the 
close involvement of all relevant stakeholders in all phases and aspects of the 
research process.

The original stimulus for the establishment of TLRP was the trenchant criticism 
directed at educational research in the mid-1990s (e.g. Tooley & Darby, 1998). 
Typically these criticisms highlighted as weaknesses the following features, as 

− Lack of rigor
− Failure to produce cumulative findings
− Theoretical incoherence
− Ideological bias
− Irrelevance to schools
– Lack of involvement of teachers
– Inaccessibility and poor dissemination
– Poor cost effectiveness

Even though there were many who argued that these criticisms were not entirely justified, they represented a significant challenge to the educational research community. In any event, the initial prospectus for the TLRP could be seen as an attempt by the funding bodies and the ESRC to address these perceived weaknesses in a sustained way. Thus the stakes were high from the outset. Funders were making an unprecedented level of investment. Politicians and policy makers had high expectations in terms of relevance, impact and quality. Meanwhile the critics of educational research were watching closely. Could the research community rise to the challenge, demonstrate the value of the investment for policy and practice and at the same time avoid compromising their core values and principles as researchers?

The education research community in the UK was not alone in facing such a challenge. National education systems in all parts of the world were increasingly operating within a global political climate which emphasized performativity, standards, accountability, international competition to enhance comparative performance and the need to undergo ‘reform’. In the UK and many other countries there were strident calls for educational policy and practice to be evidence-based or at least evidence-informed (e.g. Hargreaves, 1996).

EVIDENCE-BASED POLICY AND PRACTICE, THE ROLE OF RESEARCH AND THE AIMS OF TLRP

The call for evidence-based policy and practice takes many forms, probably the most forcible coming from those who advocate the randomized control trial (RCT) as the definitive means of obtaining evidence of ‘what works’ that can with confidence be trusted by policy makers (e.g. Torgerson & Torgerson, 2001; Tymms, Merrell & Coe, 2008; and see Torgerson, Chapter 5 of this volume). This positivist approach attempts to bring to the evaluation of educational interventions the rigorous experimental control of relevant variables. The level at which randomization procedures are applied vary from the individual level, which is often hard to accommodate within conventional educational contexts like schools, to the more pragmatic ‘cluster’ RCTs, in which randomization is applied at the class or school level. Tymms, Merrell and Coe (2008) argue for the latter as having greater ecological validity, since policy decisions tend to operate at school level, and describe a study in one local authority in Scotland in which 120 schools have agreed to be randomly assigned to different forms of peer learning intervention. The appropriateness and value of RCTs can, however, be questioned on a number of grounds (see Styles, Chapter 6 of this volume). For instance, how can ethical standards of informed consent be met in the case of RCTs? How can large-scale RCTs authentically accommodate the complexity of educational contexts such as schools in which the learning is mediated by individual classroom teachers and where there so many other ‘uncontrolled’ variables operating? Among those who
There are some who would at least argue they should not be seen as a panacea (e.g. Newman, 2008).

Biesta (2007, p. 1) has pointed out that calls for evidence-based practice actually entail a “double transformation:” firstly, the transformation of educational research to enable it to provide relevant evidence; and, secondly, with the benefit of this research evidence, the transformation of educational practice into an evidence-based practice. Biesta (2007) argues that key assumptions underlying the idea of evidence-based education tend to oversimplify both the processes of learning and teaching and the complex nature of professional judgment and decision-making in education. Furthermore, according to Biesta (2007) conventional views about evidence-based practice — characterized by simple slogans like “What works?” — fail to recognize the importance of questioning the ends as well as the means of education. Expectations associated with the concept of evidence-based practice and the role of educational research to provide definitive answers are unrealistic, since research is not capable of providing definitive answers to educational questions that can be adopted “with absolute confidence either in the formulation of policy or in the pursuit of particular practices” (Oancea and Pring, 2008, p. 27). That is not to deny the importance of evidence in formulating policy or as a basis for professional decision-making, but to argue that being informed by good evidence is a necessary and desirable, but not sufficient, condition for policy and practice in education (Bridges and Watts, 2008).

Pollard (2006), TLRP Director from 2002, characterized the stance adopted by TLRP in response to these challenges as “reflexive activism” and defined this as:

a creative, but self-conscious and socially aware, form of mediation between external critique, constraint and pressure and the social practices, values and perspectives which characterize the academic field. It represents a commitment to act politically if appropriate, but also to take the initiative in the development of the research field (p. 252).

This approach affirms the goal of improving the quality of research practices and outcomes, but at the same time it associates this with a commitment to engage actively in the public realm within our democracy and, in particular, to generate independent evidence for public debate and the decision-making processes of both practitioners and policy makers (see also Pollard, 2007, p. 639). This stance presents a difficult balance in order to avoid, on the one hand, the accusation from the academic community of conformity with (or even ‘selling out’ to) the politicians and the policy makers and, on the other, the continuing criticism from funders, politicians and others of resistance to change on the part of academic researchers and the lack of relevance of research findings to policy and practice.

One of the challenges for TLRP, as a uniquely large national program, was to find a way to encompass, if not to reconcile, differing and even conflicting traditions of research within its portfolio of activities. Any move away from the strict scientific approach might have been seen as undermining the warrant for the credibility of the findings of the various projects. However, the importance of ensuring ecological validity demanded that researchers adopt approaches that
authentically reflect the nuances of individual understandings and shared meanings among those involved in educational settings – learners, teachers and other stakeholders – as well as the complexity and the situational uniqueness of the circumstances and conditions operating in any given educational context. Broadly speaking, the methodological diversity displayed by the projects that were funded by TLRP shows that the Program has encouraged multi-method approaches with fitness for purpose as guiding principle rather than allegiance to particular research traditions. Both qualitative and quantitative approaches have been supported and published research reports from TLRP funded projects often claim methodological triangulation as warrant for conclusions reached. But what was all of this research activity trying to achieve?

The broad aims of TLRP can be stated under the following headings:

**Learning**
TLRP aims to improve outcomes for learners of all ages in teaching and learning contexts across the UK.

**Outcomes**
TLRP studies a broad range of learning outcomes. These include both the acquisition of skill, understanding, knowledge and qualifications and the development of attitudes, values and identities relevant to a learning society.

**Life course**
TLRP supports research projects and related activities at many ages and stages in education, training and lifelong learning. The Program is concerned with patterns of success and difference, inclusion and exclusion through the life course.

**Enrichment**
TLRP commits to user engagement at all stages of research. The Program promotes research on teaching and learning across disciplines, methodologies and sectors, and supports various forms of national and international cooperation and comparison.

**Expertise**
TLRP works to enhance capacity for all forms of research on teaching and learning, and for research-informed policy and practice. This work is the particular focus of the Program’s research capacity building strategy.

**Improvement**
TLRP develops the knowledge base on teaching and learning and contributes to the improvement of policy and practice in the UK. The Program works to maximize the impact of its research.

In order to achieve these aims TLRP made six strategic commitments which shaped the development of the program over the best part of a decade. These were:
user engagement for relevance and quality;
knowledge generation by project teams;
knowledge synthesis through thematic activities;
knowledge transformation for impact;
capacity building for professional development; and
partnerships for sustainability.

The following sections will focus on user engagement, knowledge synthesis and knowledge transformation for impact, and attempt to illustrate how these commitments have been fulfilled with reference to selected examples drawn from the portfolio of research supported by TLRP. But first, let us turn to the nature of knowledge and epistemological issues that are raised by TLRP’s attempts to achieve its aims.

**EPISTEMOLOGICAL ISSUES IN EDUCATIONAL RESEARCH**

As can be seen from the list of strategic commitments listed above considerable emphasis was placed by TLRP on the idea of research being about creating new knowledge. Of course this raises a number of epistemological issues about the nature of evidence, precisely what kind of knowledge should we expect research to create in a field as complex as education and how such knowledge can meaningfully be said to be cumulative. The questions of what constitutes evidence and new knowledge in educational research and the nature of the relationship between research and policy were addressed from a philosophical perspective by a team led by David Bridges with TLRP support in the form of a Thematic Seminar Series (Bridges, Smeyers and Smith, 2008). For example, Bridges and Watts (2008) examined epistemological considerations in relation to educational research and policy, highlighting that policy is itself a complex concept and that, while there is a role for research in providing relevant evidence for the formulation of policy, there are inescapable normative and ethical considerations in the field of education which make the relationship between research and policy profoundly problematic.

Somekh (2007) has risen to this kind of challenge by advocating the development and use of “speculative knowledge”, which she describes as “engaged, opportunistic and political” (p. 198). Speculative knowledge, in Somekh’s terms, “creates best guesses for possible social futures on the basis of research into current social practices” (p. 204). This perspective emphasizes the complexity of the phenomena of interest to the educational researcher. Complexity theory, which emphasizes the multiplicity and interconnectedness of variables and factors operating in any given system (Radford, 2006), can be set against more reductionist, positivist perspectives traditionally associated with a scientific approach to researching phenomena. Citing the work of Cilliers on the characteristics of systems that may be defined as complex, Radford (2006) asserts that classrooms and schools can indeed be characterized as highly complex contexts in which to carry out research, that educational processes are intrinsically problematic for researchers and practitioners alike and that, therefore, the limitations both on the degree to which control over educational outcomes can in fact be exercised by practitioners and managers and on the capability of research to provide definitive advice to schools should be more readily acknowledged. From the perspective of complexity and
chaos theories, Radford (2006) argues that education “becomes a much more open enterprise with the emphasis on a more flexible, tentative, imaginative and creative response to the multiple points at which alternative eventualities become possibilities” (p. 188).

Concepts such as complexity and speculative knowledge certainly challenge conventional assumptions about the inter-relationships between research, policy and practice. Advocates of a more tentative, nuanced approach would argue that researchers should grasp the opportunities afforded by such an approach to impact on innovation and improvement in education in more imaginative and improvised ways. Somekh (2007) challenges the research community to accept that recognizing the importance of speculative knowledge involves them being more willing to take risks “by going beyond what can be said with certainty on the basis of reliable and valid warrants” (p. 205). This may, of course, be deeply uncomfortable to educational researchers (including those involved in a program as prominent as TLRP), steeped as they are in a tradition which prizes rigor and transparency. At the same time the associated lack of certainty will be potentially irritating to policy makers seeking clear and dependable advice. Nevertheless, Somekh (2007) portrays speculative knowledge as both a liberating and an empowering concept.

To expect categoric certainty as a product of social science would indeed be naïve, suggesting a lack of ontological and epistemological awareness. Therefore, no matter how challenging it may be, there is a need to promote a more realistic public perception of the status of knowledge of the sort produced by educational research, and social science generally. At the same time, we need to establish alternative and constructive (but realistic) stances. For instance, in relation to educational research Pollard (2006) argued:

As in other fields of applied social science, cause-and-effect (or ‘what works’) is often hard to establish precisely, but educational research nevertheless has the potential to provide information, analysis and insight, and thus to significantly improve understanding and decision-making by users . . . Educational problems, of practice or policy, are often complex and immediate – but are invariably grounded in more enduring issues which merit both practitioner enquiry and sustained, cumulative, multidisciplinary, social scientific research. The introduction of evidence can clarify the key factors and major dilemmas which practitioners and/or policy-makers face in decision-making (p. 254).

Such arguments do not mean, however, that education and educational researchers (and other social scientists) can avoid their responsibility to work to achieve the highest possible degrees of confidence in their findings and analyses. This remains a major challenge to research in education. Furthermore, we would argue that the status of the field is undermined when issues of research rigor, the quality of evidence and the nature of new knowledge generated are merely glossed over. It is not sufficient to argue that we should be content with speculative knowledge. TLRP in attempting to address these issues has, for example, fostered sustained debates on the nature of the ‘warrants’ that are available for project findings. As a
consequence, each project’s Research Briefing includes a statement outlining the methodological foundations of the study.

USER ENGAGEMENT FOR RELEVANCE AND QUALITY

User engagement has been a major strategic commitment of TLRP from its inception and collaborative approaches have been advocated across all phases of project activity. Involving other stakeholders in the research process may be seen as a step towards recognizing that research knowledge in the field of education needs to be more speculative, more nuanced and more open to a range of perspectives. However, research construed as collaborative enquiry, which entails engaging and sustaining the commitment of the widest possible range of partners over the long term, is very demanding on resources, energy and intellectual capital. Why should such energy, resources and intellectual capacity be invested in user engagement in research? Perhaps there are more benefits than simply facilitating better lines of communication for the dissemination of research findings, or improving the perceived warrant for such findings. The TLRP proposition has been that a commitment to more fundamental forms of user engagement actually carries the potential to enhance the quality of the research itself. The involvement of users in genuinely collaborative research activity may permit a more comprehensive and sound conceptual grasp of issues under investigation by the incorporation of a wider range of perspectives and understandings. Methodologically, it may also represent a more ecologically valid way of investigating education as an area of ‘real world’ concern. Edwards, Sebba and Rickinson (2007), drawing on a TLRP-funded seminar series on forms of user engagement, encouragingly found evidence of “reciprocity of the co-construction of research-based knowledge across organizational boundaries emerging in different ways across the TLRP projects” and that this mirrors more general arguments for the “reciprocity of professionalism” (p. 659). The co-construction of knowledge is arguably the purpose of collaborative research and enquiry and is a process which has been identified by Cassidy et al. (2008) as the aim of communities of enquiry, which they propose as a model for educational research.

TLRP projects have adopted a wide range of innovative strategies for user engagement, going beyond the involvement of stakeholders in project advisory groups and dissemination opportunities. Where the research involved some kind of educational intervention this was, more often than not, mediated in an authentic and ecologically valid way by the appropriate professional practitioners, rather than by researchers. For example, the “Supporting Group Work in Scottish Schools” project involved 29 teachers who implemented enhanced group work in their classrooms over a one-year period (Christie et al., 2009). In other school-based projects, teachers and practitioners have themselves been involved both as researchers and as full members of project teams. A second illustrative case is provided by the “Early Professional Learning” project, led by Jim McNally, in which a cadre of six practitioners were recruited as researchers. These practitioners not only acted as participant observers, conducting interviews and administering other research instruments, but were also fully involved in the work of project team
across all stages of the research including the development and refinement of research tools, data analysis and ongoing planning (McNally and Blake, 2008).

In addition to project teams working closely with practitioners and others in their research sites, TLRP fostered links with key national organizations with potentially ‘high leverage’, in relation to dissemination and impact activity both at the individual project level and at the wider Program level. The TLRP Directors’ Team maintained links with high-leverage user organizations in each educational sector and in each part of the UK. TLRP has also worked directly with UK governments to maximize the use of its research and has been represented on almost all significant national bodies for the coordination of education research over the period.

KNOWLEDGE SYNTHESIS THROUGH THEMATIC ACTIVITIES

TLRP has adopted a comprehensive conceptual framework for its thematic initiatives and development. Figure 2.1 shows this conceptual framework in schematic form. An additional function of this model is that it provides a tagged map of linked thematic resources available on the TLRP website.

![Figure 2.1. TLRP’s conceptual framework for thematic development](image)

In the present chapter, space only allows for a few of these themes to be briefly addressed.

**Learning Outcomes**

One of the first thematic activities supported by TLRP was an attempt to map and define what was meant by the term “learning outcomes.” From the inception of the Program this was seen as a highly problematic issue deeply associated with
fundamental questions about what is valued in education. Since teaching and learning, as defined by TLRP, aims to equip learners with the resources to participate in society as active citizens, to contribute to economic development and to flourish as individuals, a broad range of outcomes is appropriate, including indicators of ‘soft’ skills. Each of TLRP’s projects has worked in different ways on defining, developing and investigating learning outcomes and hence a wide range of measures have been required to obtain evidence of change in the different domains of learning under consideration – social, emotional, and cognitive. The Supporting Groupwork in Scottish Schools project, for example, reported significant learning gains attributable to the quality of collaborative dialogue during enhanced group work both in terms of science attainment and in terms of classroom social relationships (Howe et al., 2007; Tolmie et al., 2009).

Research Capacity

Another theme cutting across the whole of the TLRP portfolio of activity is research capacity building. Among the criteria against which all project proposals were evaluated was whether the proposed research could significantly contribute to capacity building. A formal and informal, needs-based, research training program was initially funded by TLRP and coordinated by the team based at the University of Cardiff. This was subsequently followed by the adoption of the social practices model of research capacity building as developed by the Applied Educational Research Scheme (AERS) in Scotland in which inexperienced researchers are provided with opportunities to join or be associated with funded project teams and hence gain real experience of all aspects of the research process. The capacity building strategy was also complemented by the development of a suite of web-based resources for research training to be made available as “public goods” in collaboration with AERS, NCRM and the British Educational Research Association, the UK national individual membership organization for researchers in education.

Evidence Informed Principles of Pedagogy

One of the main strands of synoptic work carried out by TLRP which drew from findings across all of the projects in the Program was the synthesis of ten key principles of pedagogy (See www.tlrp.org/themes/themes/tenprinciples.html#One). To illustrate, three of the ten principles are as follows:

Effective pedagogy equips learners for life in its broadest sense

Learning should aim to help individuals and groups to develop the intellectual, personal and social resources that will enable them to participate as active citizens, contribute to economic development and flourish as individuals in a diverse and changing society. This means adopting a broad conception of worthwhile learning outcomes and taking seriously issues of equity and social justice for all.
Effective pedagogy requires learning to be scaffolded

Teachers, trainers and all those, including peers, who support the learning of others, should provide activities, cultures and structures of intellectual, social and emotional support to help learners to move forward in their learning. When these supports are removed the learning needs to be secure.

Effective pedagogy fosters both individual and social processes and outcomes

Learners should be encouraged and helped to build relationships and communication with others for learning purposes, in order to assist the mutual construction of knowledge and enhance the achievements of individuals and groups. Consulting learners about their learning and giving them a voice is both an expectation and a right.

Each of the ten principles is warranted by evidence from several specific projects and the website and associated documentation allows interested parties to unpack and interrogate both the evidence base and the implications for policy and practice of each principle. In an attempt to ensure such syntheses of research evidence are as accessible and influential as possible a very wide range of impact and dissemination strategies have been adopted. These are addressed in the following section.

KNOWLEDGE TRANSFORMATION FOR IMPACT

TLRP has attempted to challenge traditional conceptions of research impact. Rather than impact being construed as a simple linear flow starting with research followed by transformation and dissemination of findings, then adoption into practice and policy, it has been seen much more as a collaborative process from the outset in which the research itself is reflexive, interactive, iterative, constructive, distributed and transformative. As an integrated program with clear intentions to add value to research activity TLRP has engaged in a much broader range of activities than just undertaking and disseminating research. For TLRP impact, therefore, includes increasing awareness of new ideas and fostering openness to change as well as direct influence on practice and policy. Working for impact is embedded in everything TLRP has attempted to do from formulating research questions through to evaluating research influence on individual practice, not just disseminating findings of research projects. It has also been recognized that impact comes not just from “successful” research, since lessons from unexpected outcomes of activities may be just as valuable.

However, effective communication in support of influence and impact has for a number of reasons been very challenging for TLRP, as for other major research initiatives in applied fields like education. The world of teaching and learning is very complex. Producing high quality research is a *sine qua non* for maximizing impact, but this is not necessarily the main consideration among policy makers and practitioners, who may be more concerned about relevance to, and congruence
with, current priorities, than what might be seen as more esoteric indicators of research quality. Generalizing from the immediate research setting across varied teaching settings is difficult for researchers, change agents and policy makers alike. Practitioners and their managers also have to consider whether the ideas and findings in question could be applied in their particular circumstances. Furthermore, policy makers, managers, teachers and other practitioners tend to look for well-packaged solutions for defined problems, while research realities are invariably more messy. They work in very different contexts and to different timescales from the research community. They are also likely to be more receptive to research if it supports a desired policy direction and offers succinct, clear-cut advice. Thus transforming research so it has wider value and can be applied by stakeholders to their own circumstances is notoriously difficult and not without dangers. On the one hand researchers must avoid making overstated claims for what might be achieved in practice. On the other hand, researchers must not oversimplify their accounts of research findings in ways that might obscure the essential complexity and problematic nature of educational practice. (See http://www.tlrp.org/manage/admin/caip.html)

In attempting to navigate these hazardous waters, TLRP has drawn on its key strengths as a well resourced, but independent Program, which strives to maintain its integrity and profile in terms of research quality and which at the same time enjoys a high reputation and level of credibility with key decision making bodies, policy makers and professional practitioners. Because of its level of funding TLRP has been able to offer a wider range of other assets beyond the research projects – research reviews, conceptual and methodological developments, policy inputs, emerging findings, capacity building activities and resources and opportunities for interaction with a wide range of social science views. TLRP has also invested effort in establishing and maintaining networks of communication and influence in organizations, which share the broad aims of the Program and want to engage with the process of putting research evidence at the forefront of the day to day operational context of teachers, trainers and other practitioners.

The Impact Strategy adopted by TLRP involved the following seven components:

– working with networks of practitioners, learners and others to deliver, transform and communicate research evidence and other assets (reviews, methodologies, perspectives, etc.) to the widest possible number of their peers, and to convince practitioners to apply its recommendations to their practice;
– engaging with a wide range of user organizations and other stakeholders to embed the aims, approaches and findings of the Program in the cultural context of both research and practitioner communities;
– contributing to (and where appropriate organizing and leading) strategic debates about teaching and learning, to influence policy and practice;
– working with key organizations in the field to raise the profile of evidence-informed teaching and learning as a significant and effective route to lasting improvements in attainment of learners;
– communicating conceptual, methodological and practical approaches at both Program and project level to research, practitioner and relevant policy communities;
supporting training and other capacity-building activity which improves the
ability of the teaching and learning community to undertake relevant research
and transform it appropriately for a range of audiences; and

- facilitating the learning and co-operation opportunities generated by discussion
between projects within the Program, to develop both conceptual and practical
advances in transforming research into impact.

TLRP has adopted a range of different modes of dissemination, both conventional
and unconventional, including a coherent package of corporately badged outputs in
the form of briefing papers from each project, conventional reports, practitioner
guides and applications, a series of academic books and also very influential themed
synoptic commentaries drawing upon and warranted by evidence from a large
number of projects. The Program website (www.tlrp.org) is itself a comprehensive
resource representing its enormous portfolio of activities – and it also serves as a
powerful portal to other sources of information. Individual project teams have been
couraged to be innovative in terms of dissemination strategy. For example, the
“Home-school knowledge exchange” project led by Martin Hughes and colleagues
at University of Bristol commissioned the production of a powerful drama
performance based on the authentic narrative voices of children and young people
reflected in the accounts they had provided to the research team of their
experiences of primary to secondary school transition.

The value added by the program-wide perspective and the collective potential of
such an extensive portfolio of research projects is demonstrated by the impact
of the synoptic TLRP Commentaries, of which there have been eleven at the time
of writing. These are attempts to respond from a research perspective to issues
considered to be of high relevance to public policy and professional practice.
Judging by the frequency of downloading from the TLRP website (over 600,000
times), these commentaries have been very widely taken up. For example, the
commentary on Neuroscience and Education (Howard-Jones, 2007) exceeded
50,000 downloads in its first month. Another very widely welcomed and highly
influential commentary relating specifically to school learning, edited by a
journalist, Diane Hofkins (2007) is: “What is and what might be? Principles into
practice: a teachers’ guide to research evidence on teaching and learning”. This
includes an accessible distillation of the ten principles for pedagogy outlined above
and provides illustrative detail of relevant research projects accompanied by
resources, including a poster and digital video recording of a series of authentic
case studies of relevant classroom practices. There are those who might argue that
such distillations of research evidence must inevitably oversimplify and even
trivialize what are essentially complex questions about learning and teaching.
However, as with the other Commentaries, the warrant for the principles presented
is clearly founded upon research evidence and any reader has the opportunity to
follow up for themselves particular lines of enquiry through links to other forms of
project output. It is also arguably an important task for the research community to
engage in the process of distillation of key principles. Setting oneself such a task
serves the researcher as a stimulus to stand back from the detail of any given piece
of work and look for patterns and insights in emerging themes which can not
only be related to established theory but also serve to build new theoretical understandings.

CONCLUSION: FOUR METAPHORS FOR RESEARCH IN EDUCATION

In concluding this chapter, we have chosen to refer to the four metaphors for educational research offered by Marilyn Cochran-Smith (2007), namely: research as weapon; research as warranty; research as foundation; and research as stance.

Research as Weapon

This metaphor recognizes that research evidence can be used to mount robust challenges to current policy and practice and that research activity itself can have an adversarial, competitive dimension as part of the culture of performativity that governs academic life. As described in the early part of this paper, TLRP has had to be ready to defend itself from possible attack from politicians, funding bodies, the policy community and the professions as well as the rest of the research community. It was greatly helped in this by the rise of the concept of ‘evidence-based policy making’ which, despite its susceptibility to rhetoric, provided a strong legitimation for the investment. The Program was also assiduous in engaging with senior civil servants in all UK government education departments and key agencies and with the Education Front Bench Teams of the three major political parties at Westminster. This was part of an explicit strategy of ‘constructive engagement’ which, in combination with Commentaries, Research Briefings, public presentations and other impact work, were designed to contribute evidence to inform both democratic debate and decision-making. Research evidence produced by TLRP has also certainly been used by politicians or policy makers in support of particular policy priorities – perhaps meriting the ‘weapon’ metaphor – but the Program has tried to ensure that its analysis has been available to all. TLRP’s own analytic conclusions, and its moral commitment to the improvement of teaching and learning, were promoted over many years and through a coherent set of initiatives. Indeed, it is not by chance that TLRP’s Teachers’ Guide to school findings is subtitled “What is and what might be?” (Hofkins, 2007). Thus, whilst TLRP has not exactly welded its research findings into a weapon, they have been knowingly accumulated in a form which has the potential to gradually undermine naive policy commitments to performativity in favor of policies and practices founded on an understanding of the nature of learning.

Research as Warranty

This metaphor depicts research as providing the warrant or justification for educational proposals and decisions. It highlights the responsibility that rests upon researchers to ensure that the conclusions they reach in reporting the findings of their research are warranted by the evidence they have gathered and the logical coherence of their reasoning. Instructions from TLRP to research teams on writing
about warrants say that warrants are “Key reasons why readers should have confidence in your conclusions.” Such reasons might include: the elegance of project design; empirical robustness; conceptual and theoretical coherence; the degree of user engagement and validation; careful/systematic review of other evidence; and peer review. TLRP Research Briefings contain a statement indicating what the warrants are for project conclusions and the content of all other TLRP publications, including its Research Commentaries on issues of relevance to policy and practice, are substantiated by careful reference to relevant research evidence.

Research as Foundation

As has been argued above, research should engage with policy and practice. However, research is much more than a matter of gathering evidence in response to the educational policy agenda or in order to evaluate some aspect of current professional practice. Research should also provide the foundation on which we build our understanding of the world. In short, it should be concerned with generating new knowledge. Whitty (2006) drew a distinction between two terms used to describe research in education by referring to “educational research” as research concerned in one way or another with improving policy and practice while the term “education research” should be used more broadly to characterize the whole field. This helps us to recognize that research in education should address fundamental questions which transcend the exigencies of the current political and policy priorities. OECD (2002) in its examiners’ report on educational research and development in England called for more research in Pasteur’s Quadrant, namely, research that both enhances our fundamental understanding of (educational) phenomena and at the same time informs practice. TLRP has very definitely attempted to foster this kind of approach, but it is perhaps too early to judge whether the eclectic portfolio of research projects it has supported has indeed informed fundamental understandings in the field of education. Work on the ten principles of effective teaching and learning (in versions for both schooling and post-compulsory education) and the production of international handbooks (for SAGE) are contributions to this goal, but only the passage of time will enable an evaluation of the impact of these suggestions to be made.

Research as Stance

In an applied field like education, perhaps above all, we should see research as a stance that could usefully be adopted by all stakeholders. Ideally researchers, policy makers and practitioners should have a “researcherly” disposition and be research-minded. In other words all stakeholders in education should adopt a questioning, enquiring approach to whatever role they are attempting to fulfill. Cochran-Smith (2003) in describing the role of research or “inquiry” in the teaching profession uses the metaphor of stance to suggest:

both orientational and positional ideas, to carry allusions to the physical placing of the body as well as to intellectual activities and perspectives over
time. In this sense, the metaphor is intended to capture the ways we stand, the ways we see, and the lenses we see through. Teaching is a complex activity that occurs within webs of social, historical, cultural and political significance. Across the life span an inquiry stance provides a kind of grounding within the changing cultures of school reform and competing political agendas (p. 8).

In this respect, we should seek to create a culture of evidence that can enhance professional practice in education, but such a culture should also be characterized by a critical activism (Sachs, 2003). This is indeed the approach which has been strongly advocated by TLRP and is reflected in many ways across its portfolio of activities. It appears that the body of work that has been produced by TLRP over the past decade is already proving influential in a wide range of professional and policy contexts – for instance, it can be explicitly tracked in the publications of all four General Teaching Councils in the UK, in the Learning and Skills Improvement Service, the Higher Education Academy and even the work of the Economic and Social Research Council. Programs of professional education in the UK increasingly aspire to foster research as a stance among the next generation of educational practitioners and have through the legacy of TLRP not only a source of research evidence but also a source of inspiration for what they might achieve in their own professional practice. More generally, it can be said that the overall trend of government policy in the countries of the UK is highly consistent with the implications of TLRP findings, as communicated to appropriate government bodies. There is no simple connection of course, but as ideas move iteratively in the ebb and flow of biography and history, TLRP can claim to have at least made a substantial contribution to thinking about education policy and practice in the first decade of the 21st century.

REFERENCES


