Shavelson, R.J. & Towne, L. (Eds.) (2002), *Scientific research in education*. National Academies Press.

Executive Summary

No one would think of getting to the Moon or of wiping out a disease without research. Likewise, one cannot expect reform efforts in education to have significant effects without research-based knowledge to guide them. Scientific research in education can shed light on the increasingly complex and performance-driven U.S. education system. Such research covers a wide range of issues, including teaching second-language learners, measurement of achievement and self-concept, the biological and psychological basis of language and cognition, public school finance, and postsecondary and life-long learning outcomes.

There is long-standing debate among scholars, policy makers, and others about the nature and value of scientific research in education and the extent to which it has produced the kind of cumulative knowledge expected of scientific endeavors. Most recently, this skepticism led to proposed legislation that defines what constitutes rigorous scientific methods for conducting education research.

That proposal, coupled with rising enthusiasm for evidence-based education policy and practice, led to this National Research Council study to examine and clarify the nature of scientific inquiry in education and how the federal government can best foster and support it. Specifically, the charge to the committee was to "... review and synthesize recent literature on the science and practice of scientific educational research and consider how to support high quality science in a federal education research agency." We did not attempt to evaluate the quality of bodies of existing research, of existing researchers in the field, or of the existing federal research function because that would have constituted a monumental challenge and we judged it to be beyond the scope of our charge. Instead, we adopted a forward-looking approach that draws on lessons from history and identifies the roles of various stakeholders (e.g., researchers, policy makers, practitioners) in fulfilling a vision for the future of education research.

NATURE OF SCIENCE

At its core, scientific inquiry is the same in all fields. Scientific research, whether in education, physics, anthropology, molecular biology, or economics, is a continual process of rigorous reasoning supported by a dynamic interplay among methods, theories, and findings. It builds understandings in the form of models or theories that can be tested.

Advances in scientific knowledge are achieved by the self-regulating norms of the scientific community over time, not, as sometimes believed, by the mechanistic application of a particular scientific method to a static set of questions.

The accumulation of scientific knowledge over time is circuitous and indirect. It often traverses highly contested territory—by researchers and other interested parties—and progresses as a result of a not-so-invisible hand of professional skepticism and criticism. Rarely does one study produce an unequivocal and durable result; multiple methods, applied over time and tied to evidentiary standards, are essential to establishing a base of scientific knowledge. Formal syntheses of research findings across studies are often necessary to discover, test, and explain the diversity of findings that characterize many fields. And it takes time to build scientific knowledge, whether in the physical, life, and social sciences or in areas related to education.

The scientific enterprise depends on a healthy community of researchers and is guided by a set of fundamental principles. These principles are not a set of rigid standards for conducting and evaluating individual studies, but rather are a set of norms enforced by the community of researchers that shape scientific understanding. We conclude that six guiding principles underlie all scientific inquiry, including education research:

SCIENTIFIC PRINCIPLE 1
Pose Significant Questions That Can Be Investigated Empirically

Moving from hunch to conceptualizing and specifying a worthwhile question is essential to scientific research. Questions are posed in an effort to fill a gap in existing knowledge or to seek new knowledge, to pursue the identification of the cause or causes of some phenomena, or to formally test a hypothesis. Ultimately, the final court of appeal for the viability of a scientific hypothesis or conjecture is its empirical adequacy. Scientists and philosophers commonly hold that the testability and refutability of scientific claims or hypotheses is an important feature of scientific investigations that is not typical in other forms of inquiry. The questions, and the designs developed to address them, must reflect a solid understanding of the relevant theoretical, methodological, and empirical work that has come before.

SCIENTIFIC PRINCIPLE 2
Link Research to Relevant Theory

It is the long-term goal of much of science to generate theories that can offer stable explanations for phenomena that generalize beyond the particular. Every scientific inquiry is linked, either implicitly or explicitly, to some overarching theory or conceptual framework that guides the entire investigation. Science generates cumulative knowledge by building on, refining, and occasionally replacing, theoretical understanding.

SCIENTIFIC PRINCIPLE 3
Use Methods That Permit Direct Investigation of the Question

Methods can only be judged in terms of their appropriateness and effectiveness in addressing a particular research question. Moreover, scientific claims are significantly strengthened when they are subject to testing by multiple methods. While appropriate methodology is important for individual studies, it also

has a larger aspect. Particular research designs and methods are suited for specific kinds of investigations and questions, but can rarely illuminate all the questions and issues in a line of inquiry. Therefore, very different methodological approaches must often be used in various parts of a series of related studies.

SCIENTIFIC PRINCIPLE 4
Provide a Coherent and Explicit Chain of Reasoning

At the core of science is inferential reasoning: explanations, conclusions, or predictions based on what is known and observed. Making scientific inferences is not accomplished by merely applying an algorithm for using accepted techniques in correct ways. Rather, it requires the development of a logical chain of reasoning from evidence to theory and back again that is coherent, shareable, and persuasive to the skeptical reader. The validity of inferences made through this process is strengthened by identifying limitations and biases, estimating uncertainty and error, and, crucially, systematically ruling out plausible counterexplanations in a rational, compelling way. Detailed descriptions of procedures and analyses are critical to permit others to critique, to analyze, and to attempt to replicate, a study.

SCIENTIFIC PRINCIPLE 5
Replicate and Generalize Across Studies

Scientific inquiry emphasizes checking and validating individual findings and results. Since all studies rely on a limited set of observations, a key question is how individual findings generalize to broader populations and settings. Ultimately, scientific knowledge advances when findings are reproduced in a range of times and places and when findings are integrated and synthesized.

SCIENTIFIC PRINCIPLE 6
Disclose Research to Encourage Professional Scrutiny and Critique

Scientific studies do not contribute to a larger body of knowledge until they are widely disseminated and subjected to professional scrutiny by peers. This ongoing, collaborative, public critique is an indication of the health of a scientific enterprise. Indeed, the objectivity of science derives from publicly enforced norms of the professional community of scientists, rather than from the character traits of any individual person or design features of any study.

APPLICATION OF THE PRINCIPLES TO EDUCATION

While all sciences share common principles, every field of study develops a specialization as the principles are applied. Education has its own set of features—not individually unique from other professional and disciplinary fields of study, but singular in their combination—that gives rise to the specialization of education research.

Education is multilayered, constantly shifting, and occurs within an interaction among institutions (e.g., schools and universities), communities, and families. It is highly value laden and involves a diverse array of people and political forces that significantly shapes its character. These features require attention to the physical, social, cultural, economic, and historical environment in the research process because these contextual factors often influence results in significant ways. Because the U.S. education system is so heterogeneous and the nature of teaching and learning so complex, attention to context is especially critical for understanding the extent to which theories and findings may generalize to other times, places, and populations.

Education research as a profession has defining features as well. For example, multiple disciplinary perspectives bear on the study of education. Furthermore, conducting education research that involves studying humans (e.g., students, teachers) is governed by the need to ensure ethical treatment of these participants. Finally, education research depends on its relationships with practice. These links exist along a spectrum: some types of research require only a weak connection; others require full partnerships with schools or other entities. In order to analyze state assessment data, parents and schools have to agree to a test administration. To study mechanisms by which interventions increase student achievement would require long-term partnerships between research and practice.

The features of education, in combination with the guiding principles of science, set the boundaries for the design of scientific education research. The design of a study does not make the study scientific. A wide variety of legitimate scientific designs are available for education research. They range from randomized experiments of voucher programs to in-depth ethnographic case studies of teachers to neurocognitive investigations of number learning using positive emission tomography brain imaging. To be scientific, the design must allow direct, empirical investigation of an important question, account for the context in which the study is carried out, align with a conceptual framework, reflect careful and thorough reasoning, and disclose results to encourage debate in the scientific community.

DESIGN PRINCIPLES FOR FOSTERING SCIENCE IN A FEDERAL EDUCATION RESEARCH AGENCY

How should a federal education research agency be designed if the goal is to foster scientific research on education, given the complexities of the practice of education, the stringencies of the scientific principles, and the wide range of legitimate research designs? To address this question, we did not conduct an evaluation of the Office of Educational Research and Improvement (OERI), the chief existing research agency in the U.S. Department of Education. Moreover, the committee was not charged with, nor did we attempt to develop, a comprehensive blueprint for federal education research agency; that work is best left to organizational design experts and the political process. Rather, the committee developed six design principles for a federal education research agency to nurture a *scientific culture* within the agency. The precise structure itself is not the critical element. The committee's review of the processes and practices across a range of federal research agencies and of the history of the education research agency in particular suggests that it is not the nuts and bolts of agency mechanics that differentiates successful agencies from unsuccessful ones; agencies are effective when their culture supports the principles of science.

To develop such a scientific culture, the agency must have an infrastructure that is insulated from political micromanagement, supported by sufficient and sustained resources, and led by staff with top scientific and management credentials who have the flexibility to make decisions and are accountable for them. Importantly, responsibility for the success of such an agency lies with all education stakeholders. The government cannot mandate a healthy federal role. In particular, the community of education researchers—as a matter of professional responsibility—must engage in its work to promote the agency's critical role in a vibrant education research enterprise. The design principles that follow elaborate these core ideas and include suggestions for supporting mechanisms.

DESIGN PRINCIPLE 1
Staff the Agency with People Skilled in Science, Leadership, and Management

The director of the agency should have demonstrated outstanding leadership capabilities and be a respected researcher in education. Research staff should hold similar qualifications, as well as be adept at writing grant announcements, engaging with the field to identify research gaps and priorities, and assembling panels of peers to perform various tasks. Qualified staff is so critical to a healthy agency that we believe without them, little else matters. Only with such staff can the norms of scientific research in education become infused into the agency.

DESIGN PRINCIPLE 2
Create Structures to Guide the Research Agenda, Inform Funding Decisions, and Monitor Work

The research agenda must be developed through a collaborative process that engages the range of stakeholders in education. An advisory board of researchers, practitioners, business people, and policy makers (perhaps modeled after the National Science Board) could work in collaboration with an agenda-setting committee. To provide additional input to the agenda-setting process, as well as to vet research proposals, peer review is the single best, although certainly not perfect, model. Standing peer-review panels, preferably with rotating terms, can learn from, and communicate to, the field and in turn be especially strong instruments for promoting scientific progress over time. The choice of peers with excellent scientific credentials and an ability to think across areas is the key to making this commonly used mechanism work, and depends critically on an ample talent pool of peers.

DESIGN PRINCIPLE 3
Insulate the Agency from Inappropriate Political Interference

The research agency must be insulated from political micromanagement, the distortion of research agendas by excessive focus on immediate problems, and the use of the agency as a tool to promote particular policies or positions. At the same time, its work should include policy research and short-term work that is responsive to current priorities and needs. Given trends in “hybrid” federal organizations that support both education research and service-oriented programs, we suggest that the research function of an agency be organizationally separate from, though intellectually linked to, an educational improvement mission to ensure that the research mission is nurtured.

DESIGN PRINCIPLE 4
Develop a Focused and Balanced Portfolio of Research That Addresses Short-, Medium-, and Long-term Issues of Importance to Policy and Practice

Short- and medium-term scientific studies are most responsive to the need for answers to questions of pressing problems of practice and policy. Long-term studies address fundamental questions by focusing on the development and testing of theoretical frameworks. All should be organized in coherent programs of related work. The portfolio should include research syntheses as well as new scientific investigations.

DESIGN PRINCIPLE 5
Adequately Fund the Agency

Estimates of the federal investment in education research have shown it to be a few tenths of one percent of the total amount spent on public elementary and secondary education each year—far less than comparable investments for agriculture and medicine. The research budget of the OERI (and its predecessor agency, the National Institute of Education) has fallen drastically since its inception: in 1973, its budget was over $525 million; today, it is approximately $130 million (both in 2000 dollars). As funding plummeted, there has been no commensurate change in the scope of its agenda, and thus there have been few opportunities for long-term research programs. We echo the calls of several previous studies and commissions for a significantly increased research budget if its agenda is to cover the breadth of content required of its predecessors. Stagnant funding, an inconsistent commitment, or both, means that scientific research in education is not being taken seriously.

DESIGN PRINCIPLE 6
Invest in Research Infrastructure

The agency should consistently invest in infrastructure-building programs to foster a scientifically competent, highly qualified community of education researchers and to strengthen its own capacity in turn. Since an agency in many ways is a reflection of the field it supports, such programs should include investment in human capital (e.g., research training and fellowship support). Promoting ethical access to research subjects and data should be an essential task as well. An agency should also do its part to facilitate relationships between practitioners and researchers both for basic access to data as well as, in many field-based research efforts, for long-term partnerships with practitioner communities to improve the research as well as its utilization.