NORC at the UNIVERSITY of CHICAGO

NORC WORKING PAPER SERIES

Assessing Returns on Investments in Education Research: Prospects & Limitations of Using 'Use' as a Criterion for Judging Value*

WP-2017-002 | MAY 10, 2017

PRESENTED BY:
NORC at the University of Chicago
Sarah-Kay McDonald, Principal Research Scientist
55 East Monroe Street
20th Floor
Chicago, IL 60603
Ph. (312) 759-4000

AUTHOR:
Sarah-Kay McDonald

AUTHOR INFORMATION

Sarah-Kay McDonald
sarahkay58@gmail.com

* The author is grateful to: NORC at the University of Chicago for supporting the preparation of this working paper (including comments of anonymous reviewers); those who provided feedback on an earlier version of this paper presented at the November 2015 annual meeting of the American Evaluation Association; Norman Bradburn for his input at various stages in an ongoing program of research on estimating returns on research investments; and Siobhan O'Muircheartaigh for her helpful edits and suggestions. Any errors, omissions, or misinterpretations are of course the author's alone.

Table of Contents

Table of Exhibits

This paper is the product of a larger effort to identify metrics, measures, and methods for assessing returns on investments in education research. Taking current practice as a starting point, it explores the factors that have encouraged—and the limitations of—the notion that 'use' is (or should be) a key criterion for judging the value of investments in education research.

The paper begins by briefly rehearsing the 'ground practice in robust research evidence' logic that provided the underpinning for many millennial federal education reform initiatives. Next it describes a three-part approach that was expected to promote research use for educational improvement—by ensuring the rigor and relevance of research studies, and the accessibility of their findings. The third section of the paper sets these expectations in the broader context of an international culture of evidence-based decision-making that encouraged policymakers in many countries to approach the resolution of social problems in multiple domains by supporting the generation and/or requiring the application of research evidence. The fourth section argues that together these emphases on evidence-based decision-making and practice created a perfect storm—fueling presumptions regarding the adequacy of 'use' as a criterion for assessing the value of educational research, and judgments regarding appropriate metrics for assessing both individual studies and entire portfolios of education research investments. Such presumptions are not, however, supported by the many literatures addressing factors that influence individuals' decision-making and organizational behavior. Relevant insights from several of these are rehearsed in section five. Section six moves on to consider key implications for the development of program logic models, the feasibility of employing 'use' as a criterion for evaluating education research. The final section describes five important roles program evaluators can play in reshaping expectations regarding the most appropriate approaches and metrics for gauging returns on investments in education research.

Millennial Federal Education Reform

Educational Research as a Driver of Educational Reform and Improvement

The bright promise of a new millennium and decades of federal-level reform efforts notwithstanding, as the year 2000 approached the American public education system was found sorely wanting. Its students' performance overall was judged somewhere between lackluster and unacceptable on a wide range of outcomes, from assessments of student learning via standardized tests of basic skills, to graduation rates and matriculation to courses of study at institutions for higher education. Shortcomings and disparities in educational attainment in the physical and biological sciences, information and computing technologies, engineering, and mathematics were (and remain) particularly troubling, as educational and learning outcomes in STEM continue to be singled out as critical contributors to national economic comparative and competitive advantage.[1]

Against this backdrop, federal policymakers were urgently seeking solutions to the educational challenges and inequalities facing U.S. elementary and secondary school students. As the 20th century drew to a close, these solutions were increasingly sought from individuals outside the local, state, and federal educational systems. Earlier a National Research Council (NRC) Committee on the Federal Role in Education Research had concluded the (now defunct) Office of Educational Research and Improvement (OERI) should encourage research utilization to improve educational outcomes (see NRC, 1992). Increasingly, policymakers turned to scholar-experts in institutions of higher education, nonprofit organizations, and think tanks to provide not just third-party perspectives, but also dispassionate data, emerging evidence, and innovative insights. Educational researchers were asked (and funded) to document the sources of, and suggest new approaches to resolving, persistent educational problems.

Significantly, legislation began to underscore the importance attached to research as a tool for improving teaching and learning. While ultimately not enacted into law, Rep. Castle's bill to reauthorize OERI[2] was pivotal for its positioning of research as providing "the basis for improving academic instruction and learning in the classroom" (p. 25), a posture that would become increasingly familiar. The following year, H.R. 1 in the 107th Congress similarly positioned educational research as a critical driver of educational reform and improvement. Enacted as the *No Child Left Behind Act of 2001* (PL 107-110), this legislation required efforts to improve basic local educational agencies' programs explicitly to take such research findings into account. Local areas with the greatest needs were only eligible for Title 1 funding if districts could demonstrate that technical assistance, professional development, additional instructional staff, or other school improvement strategies were grounded in research evidence that clearly demonstrated a "proven record of effectiveness" (U.S. Department of Education, Office of the Under Secretary, 2002: 13). Similarly, at the federal level, NCLB prohibited the Secretary of Education from approving grants, contracts, or cooperative agreements intended to improve educational opportunities for specific populations if they were not "based on relevant research findings" (Sec.7144). Indeed, NCLB's commitment to focusing educational dollars "on proven, research-based approaches that will most help children to learn" (White House, 2002)—more colloquially, "funding for what works" (U.S. House of Representatives, Committee on Education and the Workforce, 2002)—was one of four hallmark features of the legislation singled out at the time of signing. As characterized by Feuer, Towne, and Shavelson (2002: 4), millennial era legislation "exalt[ed] scientific evidence as the key driver of education policy and practice".

Critically, federal policy not only required that reforms be research-based, it put forward a clear, three-part plan of action, with a user-focus, to promote research use for educational improvement. Part one of the plan was designed to ensure districts, states, educators, students, parents, and other caregivers could be confident consumers of education research. Clear guidelines were articulated for what constituted, and how to conduct, investigations capable of generating evidence of high technical quality—to ensure research would be rigorous. Part two of the plan was designed to assure research would also be relevant—engineered, by design, to be user-friendly, with research agendas reflecting users' needs, concerns, and priorities. Finally, part three of the plan involved making research findings accessible—'user-friendly' with respect to not just content, but also to users' informational needs and information-seeking behaviors. Public policy initiatives to implement this plan critically shaped the expectation that education research could, would, and should be used to achieve sought education and learning impacts, in the elementary and secondary grades and beyond—with important implications for conceptions of the value of educational research.

Great Expectations, Part 1

Rigor + relevance + access ⇨ use ⇨ impact

NCLB was not the first nor would it be the last federal initiative to encourage the use of research. It did, however, mark a watershed in federal efforts linking quality and accessibility of research (or the lack thereof) to its utilization—with important implications for the presumed viability of 'use' as an indicator of the value of education research.

Use and the Rigor of Research

In his last biennial report to Congress as Director of the U.S. Department of Education's Institute of Education Sciences (IES), Grover J. ("Russ") Whitehurst reminded federal legislators that forty years previously "there was no evidence that anything worked in education" (IES, 2008: 1). Whitehurst made this observation citing a 1972 report prepared by The RAND Corporation that concluded: "…research has found nothing that <u>consistently</u> and <u>unambiguously</u> makes a difference in students' outcomes" (Averch et al, 1972: x; emphasis in the original).[3] The qualifiers were important ones—the authors underscored (literally and figuratively) at several junctures that they were "<u>not suggesting that nothing makes a difference, or that nothing 'works'</u>," (Averch et al., 1972: x; emphasis in the original). Indeed their review of the literature found "numerous examples of educational practices that seem to have significantly affected students' outcomes" (*Ibid*., p. 154). Worryingly, however, the authors also found

> "…there are invariably other studies, similar in approach and method that find the same educational practice to be ineffective. And we have no clear idea of why a practice that seems to be effective in one case is apparently ineffective in another" (Averch et al., 1972: 154-155).

A question many posed regarding such apparently conflicting findings was whether these were indicative of an underlying problem with the quality of the research. Numerous individuals and organizations took up the challenge of defining what constitutes 'high-quality' education research, and how best to support its creation. In 2000, H.R. 4875 addressed the quality issue by calling upon federal executive bodies to promote, provide funds to support, offer technical assistance predicated on, and disseminate 'scientifically valid' research—defined as encompassing "applied research, basic research, and field-initiated research whose rationale, design, and interpretation is soundly developed in terms of established scientific research and that is conducted in accordance with scientifically based quantitative research standards and qualitative research standards" (p. 3). A year earlier, the *Omnibus Consolidated and Emergency Supplemental Appropriations Act of 1999* (Pub.L. 105-277) had similarly specified standards of evidence that evaluations would need to produce in order to warrant the adoption of specific educational practices (McDonald, 2009). Shortly thereafter (in autumn 2000), the U.S. Department of Education's National Educational Research Policy and Priorities Board (NERPPB) asked the NRC to "consider how to support high quality science in a federal education research agency" (NRC, 2002: 22). The Committee on Scientific Principles for Education Research Results was established in 2001 and results of the Committee's work were published in the pivotal volume *Scientific Research in Education* in 2002, the same year NCLB was signed.

In short order, NCLB was widely cited for its numerous references to requirements regarding the use, not just of education research, but specifically of "scientifically based research". Local educational agencies could receive grants conditional on assurances that their plans would "take into account . . . the findings of . . . scientifically based research". Reform strategies and targeted assistance school programs were required to use, "effective methods and instructional strategies that are based on scientifically based research". Technical assistance provided by local education agencies to schools identified for school improvement, school library media programs, comprehensive school reforms, requests for mathematics and science partnership grants, programs to promote safe and drug-free schools—all were expected to be grounded in scientifically based research.

The Act defined such research as that which "involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs". For research to qualify as 'scientifically based' it needed to meet a number of criteria, including that it employs "experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random-assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition controls…."[4] Educational practice was to be grounded, then, in education research that met specific quality standards with respect to design and analytic rigor—with a particular emphasis on internal validity. Robust evidence should stand behind claims that something would 'work' to achieve desired educational outcomes.

In an effort to increase the prospects that practitioners would have access to research-based reforms meeting these standards, the *Education Sciences Reform Act of 2002* (Pub.L. 107-279) required the Commissioner of a National Center for Education Research, established within a new U.S. Department of Education Institute of Education Sciences, to "ensure that all research conducted under [its] direction . . . follows scientifically based research standards". As investigators moved onward from exploring ideas at the proof-of-concept stage, they were expected to conduct their research pursuing ever more stringent internal validity standards. Conjectures and initial evidence suggesting promising new approaches and solutions might be evidence enough to encourage further exploration, but they would be insufficient to warrant widespread use.

In the decade that followed, considerable resources, and numerous incentives and requirements, were put in place to assist education researchers in achieving the now pervasive expectations regarding what constitutes appropriately rigorous education research methods, designs, and analysis plans. For example, federal funds were provided to support the establishment of pre-doctoral and post-doctoral training programs designed to build capacity to conduct rigorous education research. The National Science Foundation (NSF) and the U.S. Department of Education's Institute of Education Sciences (IES) collaborated on the development of a set of *Common Guidelines for Education Research and Development* designed to improve "the quality, coherence, and pace of knowledge development in science, technology, engineering and mathematics (STEM) education" by describing, and providing "basic guidance about the purpose, justification, design features, and expected outcomes from" six research genres (NSF and IES, 2013: 4).

Today, funders' program materials often provide prospective investigators with direction, guidelines, or requirements for framing research questions and designing studies. These may range from general statements regarding the importance of being "cognizant of the particular research goal of your project and how this affects the type and use of your findings" to goal-specific expectations regarding theories of action to be articulated, evidence to be provided of factors associated with outcomes of interest, including potentially mediating and moderating factors, and central design features. Such guidance is not always program-specific. Instead, program materials may reference more general (e.g., cross-program and/or interagency) expectations for what constitutes research sufficiently rigorous that practitioners and other stakeholders might be encouraged to employ it in their decision making and practice.[5] A picture begins to emerge, then, of a combination of program-specific and other supports. Such initiatives were intended to reduce the probability that research would not be 'used' because of questions regarding its scientific quality. States, districts, and local education authorities were expected to implement interventions with a high probability of achieving meaningfully positive results, and education researchers were expected to provide the robust evidence that would warrant these reforms.

Use and the Relevance of Research

While Whitehurst's tenure as inaugural Director of the Institute of Education Sciences may be particularly well remembered for its contributions to efforts to enhance the rigor of education research studies, rigor was not the only criterion IES took into account under his watch in identifying 'high quality' education research. Tellingly, Whitehurst's final biennial report to Congress in that capacity carried the title *Rigor and Relevance Redux*. In it, he stated:

> It is easy to be relevant without being rigorous. It is easy to be rigorous without being relevant. It is hard to be both rigorous and relevant, but that is the path of progress and the path taken by IES (Whitehurst, 2008: 6).

The second IES Director, John Q. Easton, continued the Institute's joint focus rigor and relevance, arguably with an increased focus on the latter. The first biennial report to Congress covering a full two-year span during which Easton served in this capacity observed:

> In his first years at IES, Easton committed to maintaining the agency's rigorous standards for research, evaluation, statistics, and assessment. Through new research priorities approved by the National Board for Education Sciences and numerous public statements, Easton also promised a renewed emphasis on relevance, accessibility, and timeliness in IES work (U.S. Department of Education, 2013: 2).

It also observed that "[u]ltimately, IES will be judged by practitioners and policymakers seeking guidance from the research community on how useful and helpful they find IES activities, studies, and products" (*Ibid.*, p. 2).

Like rigor, the emphasis on relevance was anticipated in the law. NCLB defined 'scientifically based research' as involving "the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge ***relevant*** to education activities and programs" (emphasis added). IES-funded pre- and post-doctoral training programs similarly supported the training of scholars prepared to conduct not only rigorous but also relevant education research "addressing issues that are important to education leaders and practitioners" (U.S. Department of Education, 2015a). Yes, high-quality (potentially actionable, useful, valuable) research should be rigorous; research evidence of questionable scientific, technical quality was by definition understood to be less useful than more robust findings.[6] It should also, however, be cognizant of—and appropriately address—user concerns. Relevance, then, not only went hand-in-hand with rigor, it was inseparable from it as a criterion for assessing the quality of education research.

Importantly, researchers were expected to take user needs, issues, and interests into account not only at the 'distribution of findings' stage but also in developing their investigations. Taking a page from marketing strategy, investigators were encouraged to establish the demand for the knowledge and interventions they were seeking support to produce, and then supply to meet demand. They should 'make what they can sell' (generate evidence on the topics of most interest to prospective users) rather than attempt to 'sell' (e.g., encourage interest in and use of) the evidence they wanted to generate. Accordingly, scholars were encouraged to consult and collaborate with prospective users of their findings, co-constructing research agendas, interventions, and assessments.

Work to understand and foster research-practice partnerships flourished, both with and independently from supportive federal initiatives. Bryk and colleagues advanced a science of improvement research built on "the idea of a networked improvement community that creates the purposeful collective action needed to solve complex educational problems".[7] The William T. Grant Foundation convened a national network of research-practice partnerships (RPPs) with the goal of fostering "two-way streets for learning between researchers and practitioners"[8] and the Spencer Foundation initiated a grant program intended to strengthen and provide intermediate-term stability to existing RPPs.[9] On the federal side, U.S. Department of Education and National Science Foundation programs similarly called upon prospective investigators to connect their proposed projects to practice. In notices of funding opportunities applicants were asked to explain how the results of their work would "be important . . . to education practice and education stakeholders (e.g., practitioners and policymakers)" (U.S. Department of Education, 2015b: 39), or "how products or findings might ultimately be implemented in schools, even if in the long term" (NSF, 2015: 8). Issues of collaboration were also addressed at the individual award level. For example, in 2012 an NSF-supported Research + Practice Collaboratory was established to make improvement efforts "more practical, relevant, and sustainable" by building and supporting "research-practice partnerships between educators and researchers where problems are framed and improvement efforts are designed within the existing ecology of practice", with educators "who deeply live and understand the local context . . . at the center of framing the issue and designing the solution".[10] Two years later, IES established a National Center for Research in Policy and Practice that, among other things, would undertake a descriptive study to examine three types of "purposeful attempts to increase research use by promoting greater interaction between researchers and practitioners": "research alliances; design research partnerships; and networked improvement communities."[11]

It is interesting to speculate on how the emphasis on relevance may have informed the press for scalable interventions. If relevance is understood to be a key characteristic of high-quality research, it is not unreasonable to perceive research relevant to *large* numbers of individuals and organizations as more useful than research relevant to smaller numbers (at least on that 'prospective reach' dimension). Such an expectation—alone or when combined with a press for cost-effective solutions to education and learning challenges—may have fed a perception that some of the most 'relevant', useful, valuable research was that with the greatest potential for generalizability. Concerns with the extendibility of project findings and associated potential returns on research investments may have similarly encouraged investigators to explore the feasibility of replicating positive outcomes among broader groups of learners in multiple contexts. Continuing the marketing analogy, a variety of factors pointed to the merits of pursuing a 'mass marketing' strategy—developing interventions and/or broadcasting messages about 'what worked' that were relevant to and had the potential to reach the broadest possible audience. Ideally, education researchers would innovate and design interventions for scale-up.

Prospects for scaling were soon recognized to be limited, on several dimensions.[12] Ideas and innovative approaches to resolving entrenched education challenges might be scalable, but often only with adaptation to context, population, and situation. The extent to which research findings might be 'used', then, was accordingly expected to be variable across audiences and contexts, and relevance was recognized to be relative. Accordingly, education researchers were absolved from an expectation they would identify 'what works' in absolute terms. Instead they were presented with the arguably more reasonable (yet ultimately often more challenging) task of determining what works for whom, when, and under what circumstances. Consistent with resource considerations central to the larger accountability movement, education researchers were expected, where possible, to continue to seek efficiencies through economies of scale. Crucially, in those instances in which scale-up was not warranted, investigators were expected to communicate carefully the reasons why this was (or might be) so, enabling prospective consumers and users of their results to judge relevance with reference to the specific audience(s) with whom and/or conditions within which they worked. Collectively, education researchers were expected to conduct and report their work so that the rigorous, scientifically-based findings they generated would accumulate in a vast storehouse of knowledge that could inform action. Ideally, the knowledge framework knit together from these individual study findings would identify both scalable approaches to reform and specific interventions proven effective in ameliorating particular educational inequities and resolving given learning challenges. When accessed, digested, and acted upon, it would map the pathway to reform.

Use and Accessibility of Research

The millennial wave of federal education reform, with its sustained support for careful attention to the factors influencing rigor and relevance, has had a considerable effect on the quality of education research. Yet relevance and rigor alone are widely acknowledged to be insufficient for the products of education research to be used. To facilitate use, research findings must be shared—and shared appropriately. Research providers were thus exhorted to consider carefully the implications of their findings and assume greater responsibility for making them not only more relevant but also more accessible to targeted stakeholder communities, in hopes ultimately of making them more influential (see e.g., Beyer, 1997). They were encouraged to see themselves as "instruments of social problem solving" (Lindblom & Cohen, 1979: vii), and to explore how they could become more effective in this role. They were prompted to assume at least some responsibility for the oft-lamented disjuncture between research and policymaking or practice, and pressed to disseminate research evidence effectively (e.g., in audience-appropriate terminology, via the correct channels, with the research consumer's informational needs in mind). In short, they were exhorted to take proactive steps that could keep them from falling "victim to the problems, well-rehearsed in the literature, of poor dissemination" (Kirst, 2000: 379).

Communication strategies, dissemination activities, and data access and data management plans were seen as critical to ensuring that the latent potential of research evidence was (via use) achieved. At the individual project level, educators were expected to plan these strategies at the same time as they were developing the questions they proposed to answer and the methods they would employ to do so. Message development was to commence at the time research projects were conceived, with users' interests shaping research agendas. Investigators were thus encouraged to "develop partnerships with education stakeholder groups" as a means of advancing both "the relevance of their work and the accessibility and usability of their findings for the day-to-day work of education practitioners and policymakers" (U.S. Department of Education, 2015b: 1). Doing so might involve: identifying the audiences most likely to benefit from the research; discussing the mechanisms and medium to be employed in reaching these audiences; describing the investigating team's capacity to disseminate research findings; specifying the resources required to disseminate the results; and/or developing plans for sharing the products of research.[13] Not surprisingly, such proactive efforts to encourage practitioners and policymakers to access and utilize their findings were often intertwined with other steps to enhance the perceived relevance of their work. Just as rigor and relevance go hand-in-hand in shaping perceptions of the quality of education research, relevance and access went hand-in-hand in shaping perceptions that robust findings constitute usable knowledge.

Millennial reform initiatives made numerous supports available to aid investigators in assuring the accessibility of research evidence. Multiple initiatives were undertaken to support the accumulation, synthesis, and broader distribution of funded projects' work, at the program level and across programs. Once again, the legislative proposal for the Scientifically Based Education Research, Statistics, Evaluation, and Information Act of 2000 set the tone. A National Education Library and Clearinghouse Office proposed in the legislation was charged with disseminating 'scientifically valid' research "in forms that are understandable, easily accessible and usable or adaptable for use in the improvement of educational practice by teachers, administrators, librarians, other practitioners, researchers, policymakers, and the public" (pp. 9-10). Such roles were ultimately assumed by entities including the U.S. Department of Education-funded Regional Educational Laboratories and Comprehensive Centers. Perhaps most familiar to U.S. policymakers, the IES-supported What Works Clearinghouse (WWC) was a key platform in federal efforts to increase access to robust research findings. Characterized as "the linchpin for the entire enterprise of evidence-based decisions in education" (Whitehurst, 2008: 16), the WWC identified studies that do (and do not) offer "credible and reliable evidence of the effectiveness of a given practice, program, or policy" (http://ies.ed.gov/ncee/wwc/aboutus.aspx). It gave access to information on studies that meet its standards of evidence in a variety of forms (e.g., intervention reports, practice guides, 'quick reviews', and single study reviews) accessible online (e.g., using the 'Find What Works!' interface).[14] A Doing What Works (DWW) library sought to provide the same content in a format developed with educators' evidentiary needs and interests in mind.[15] At the same time, additional supports (and requirements) were provided to assist investigators in developing appropriate plans for data sharing, ensuring public access to projects' products, and communicating research findings effectively (e.g., through online compendia of resources, training, technical assistance, professional development, and one-on-one consultancy services). In a relatively short space of time, then, investigators were encouraged to assume greater responsibility for making the results of their work accessible beyond the academy, and a variety of resources were established to ease prospective target users' access to rigorous, reliable, relevant education research findings.

Great Expectations, Part 2

Evidence Should Inform Decision-Making, and Science Policymaking Should Be 'Scientific'

If the emphasis on use of research findings to improve practice (and policy) had been unique to education, or even slightly novel, the notion that utilization signaled the value of research trajectories (or discrete research findings) might have been more skeptically received. The idea that research evidence should inform practice was, however, far from unique. Instead a broader, international culture of evidence-based decision-making encouraged education policymakers to focus their attention on the generation and use of research evidence.

The histories of evidence-based practice and decision-making in multiple fields are well documented.[16] Originally associated most closely with medicine (following Archie Cochrane's 1972 report on the importance of randomized controlled trials in determining and forecasting treatment effectiveness in the UK National Health Service), the idea that practices should be rooted in robust evidence (e.g., as opposed to anecdotal or personal experience) soon spread. In what some have characterized as a culture of 'evidence-based everything' (see e.g., Fowler, 1997; Oakley, 2002), an initial focus on the rationale for medical practices morphed into a much broader preoccupation with the extent to which health care, criminal justice, social work, educational and other policies and practices should be rooted in evidence of the effectiveness of various interventions. As in education, the call for evidence-based practices and decision-making was predicated upon the quality of the evidence—with randomized controlled trials (RCTs) typically privileged for the confidence they instill in the internal validity of investigations and any causal inferences based on study findings. Solid evidence would lead to better outcomes; better outcomes would be predicated on sound evidence.

Echoing this notion was a millennial era call for science policymaking writ large to become more evidence-based—more 'scientific'—in nature. In 2005, John Marburger (then Director of the U.S. Office of Science and Technology Policy [OSTP], and science advisor to the president) expressed dismay with the lack of evidence on impacts and returns on investments available to science policymakers (i.e., those charged with making crucial allocations of funds to achieve the nation's science and technology objectives). In a speech squarely focused on budgets, Marburger called for the "nascent field of the social science of science policy . . . to grow up, and quickly, to provide a basis for understanding the enormously complex dynamic of today's global, technology-based society" (Marburger, 2005a).[17] The premise was that generation science policy would be improved if rooted in (better) evidence on the impacts of prior science policy investments. Notable U.S. responses to Marburger's call included the establishment of a Science of Science and Innovation Policy (SciSIP) program and efforts to support the development of science metrics. A long-term aim of the former effort, spearheaded by NSF, was to create "a cadre of scholars who can provide science policy makers with the kinds of data, analyses, and advice that economics . . . provides for setting fiscal and monetary policy" (NSF, 2006: 155).[18] The latter initiative sought to provide data and tools to facilitate the assessment of R&D investment impacts.[19] As with the larger evidence-based decision-making movement, the typically quite explicit assumption behind such efforts was that with enhanced evidence of the impacts of past investments, decision makers would be in an improved position to make 'better'—more informed—decisions on new ones, or at least have a sounder basis for weighing opportunity costs and gauging prospects for success.

A Perfect Storm

'Use' as a Criterion for Assessing the Value of (Education) Research

By the turn of the century, the international culture of 'evidence-based everything' provided a ripe climate for policymakers' interests in research utilization to flourish. In this environment, millennial education reforms fostered an expectation that rigorous, relevant, accessible research findings stood a better chance of being employed to resolve persistent challenges to improving student learning and equality of educational opportunity than less robust, less relevant, less accessible evidence. At the same time, accountability concerns and a broader 'impact agenda' (Pettigrew, 2011) underscored the importance of being able to point to the effects of research investments.[20]

Against this backdrop, efforts to demonstrate prudent distribution and sound management of education research funds increasingly sought evidence that the knowledge, innovations, interventions, resources, models, and tools research projects generated were being utilized—and to good effect. Funder and evaluator initiatives to gauge the impacts of research investments using bibliometric, scientometric, and/or altmetric analyses contributed to a perception that research yielding products that could be seen to be used was more impactful and valuable than research whose findings were not so easily traced and tracked to the discovery of new interventions, policymakers' actions, or education practices. Efforts to design for use became pervasive, and evidence that research findings were being used by stakeholders beyond the academy was particularly prized. Finally, the effort to make science policymaking more scientific meant just as researchers were expected to provide 'usable knowledge' to inform education policies and practices, education research funders were expected to provide compelling evidence of the impacts of their expenditures—not only to justify previous but also to warrant future investments.

Together, this confluence of events fueled presumptions regarding the adequacy of 'use' as a criterion for assessing the value of education research. It also shaped judgments regarding appropriate metrics for assessing the contributions and impacts of both individual studies and entire portfolios of education research investments. If access were ensured, it seemed a small step to expect that educational improvements would be achieved if reforms were rooted in high-quality research. Unfortunately, it was all too tempting to presume the converse. If rigorous research led to practical action, should it not be the case that 'unraveling the ball of string' from successful reform would lead back to robust research—that use would imply quality? If research evidence were relevant, would it not be used by key stakeholders (teachers, principals, parents, district administrators, state education agencies) as they selected curricula to implement, devised school improvement plans, decided when and how to exercise school choice options, purchased teacher professional development services, established high school graduation requirements, or made any other of the host of decisions facing those who seek to provide a high-quality education for all the nation's students? If so, could indicators of use not also serve as proxy indicators of the 'relevance' dimension of quality? If the answer to these questions was 'yes', use would perhaps be a reasonable criterion to employ in judging the quality, thus assessing the value of investments in education research.

Challenges Employing 'Use' As a Criterion for Evaluating Education Research Investments

Insights from the Literature

The notion that one can infer from the use of research evidence that the underlying science is valuable—or from the absence of use that the underlying research is *not* valuable—is refuted many times over by extensive bodies of research (as well as most individuals' practical experience). Decision-making processes and the factors influencing them are simply too complex to support such straightforward conclusions.

Multiple disciplinary traditions and literatures address issues critical to modeling, supporting, and leveraging the influence of robust research knowledge on practitioners' and policymakers' thoughts, beliefs, opinions, judgments, decisions, and actions. Scholars of judgment and decision making have explored from psychological, economic, human decision analysis, organizational behavior, and other perspectives the many factors that facilitate, impede, and constrain the processes by which conclusions are reached, courses of action are charted, problems are solved, and decisions are made.[21] Such efforts to map decision making underscore the multiplicity of factors influencing individuals' thought processes, from perceptions of risk, preferences, approaches to problem framing, abilities to process information, and propensities to analysis, to experience, local knowledge, political and pragmatic factors, intuition, 'gut check', custom, personal beliefs and values, and emotion.

Closely related bodies of research focus less on decision making writ large, more on the specific roles of evidence in the decision-making practices characteristic of particular fields of practice—and practitioners more generally.[22] For example, the ways in and extent to which research is used by managers, medical practitioners, and government officials are particularly well documented—as are the prospects of encouraging such use.[23] Other studies focus on the traction achieved by particular types of scholarship (e.g., the utilization of research produced by specific domains and subfields); examples here include studies of the uses and impacts of social science[24] and management research in the UK.[25] Education research has been—and continues to be—the subject of both types of study.

Much has been written about the ways in which education research is used, and factors influencing its use beyond the academy. Issues of interest include: the process of education change; the development of knowledge utilization systems; the problems encountered in applying new knowledge; the use of research evidence in instructional improvement; and the multiple factors influencing district administrators' and classroom teachers' use and interpretation of research evidence.[26] Additional findings are eagerly anticipated, including those from the William T. Grant Foundation's program of investments on improving the use of research evidence,[27] as well as those from two Research and Development Centers on Knowledge Utilization established by the U.S. Department of Education. These Centers are expected to "develop tools for observing and measuring research use in schools, illuminate the conditions under which practitioners use research and factors that promote or inhibit research use in schools, and identify strategies that make research more meaningful to and impactful on education practice" (U.S. Department of Education, 2014: 7).[28] Whether such initiatives will dramatically change current understandings remains to be seen. For now, a common theme running through extant studies of knowledge utilization—in education and in other domains—is that multiple factors inform deciders' beliefs and values, thoughts and actions, judgments and decisions, and the relationships among them can be quite complex. Research evidence is just one—and often not a particularly prominent—entry on the list.

Studies of judgment, decision making, and research utilization also remind us that just as there are many factors influencing use; there are many different types of use to which research findings might be put.

Forward-looking statements of intent captured in funder documents directed to the investigator community (e.g., program synopses, announcements, solicitations, and requests for proposals) note the wide range of impacts education research investments are expected to have—and the many ways in which research findings would need to be 'used' to achieve them. A sample of documents issued in fiscal year 2015 describing federal mathematics and science research funding opportunities[29] amply illustrates the wide range of uses anticipated. It was expected results of this research would be employed (among other things) to: 1) shed light on persistent problems; 2) help build theory; 3) inform the design of education interventions; and 4) produce "an array of tools and strategies (e.g., curricula, programs, assessments) . . . for improving or assessing . . . learning and achievement." Interestingly, the potential uses of research findings were important considerations in even the most 'basic' STEM education research programs. For example, it was expected research evidence would be employed to "develop understandings", "catalyze new approaches", and "build and expand a coherent and deep scientific research base" on the affective dimensions of learning—deepening understandings of "what motivates and sustains learner interest . . . and what fosters engagement and persistence". It would inform both the adoption of promising interventions by districts, schools, and classroom instructors, and subsequent iterative development, refinements, and improvements upon these interventions by education researchers. Study findings would yield "assessments that . . . help teachers provide guidance to students and inform teacher decision-making; [and] provide teachers with dynamic diagnostic information about student learning in real-time" and they would influence the priority accorded particular education practices and policies on stakeholders' agendas. In sum, funders expected education research evidence would be 'used' to: define issues as "problems" meriting public (or other) solutions; secure the place of these problems on stakeholders' agendas; devise approaches and interventions to address them; identify conditions under which they should be adopted or adapted prior to implementation; and gauge their (intended and unintended) consequences.

Several scholars have suggested typologies for categorizing such uses of research. Carol Weiss—whose work is widely regarded as central to the establishment of research utilization as a field of study—identified seven models of use of social science research by public policymakers: knowledge-driven, problem-solving, interactive, political, tactical, enlightenment, and research as part of a more general intellectual enterprise of society (see Weiss, 1979).[30] Pelz (1978) distinguished instrumental, conceptual, and symbolic uses of research findings. Tseng (2012) references rational instrumental uses; tactical, symbolic, political uses; conceptual, enlightenment uses; imposed uses; and process uses in which practitioners employ knowledge gleaned from participation in (rather than findings from) research investigations.

From an evaluative standpoint, more important than the typologies employed to categorize research utilization is that *each* of the many (intended and unintended) ways in which research might be used is considered. This applies whether the purpose of the evaluation is to explore the utility of evidence for specific target audiences, or to assess the "usefulness of the social science research that government funds support" (Weiss, 1979: 426). From a practical perspective, this means efforts to ascribe value to research investments from a utilization perspective must be holistic, e.g., going beyond documenting scholarship's use by the academy (e.g., as evidenced by bibliometrics or other measures). Evaluations that fail to consider, for example, whether, by whom, and how often (single studies' or trajectories of) research findings are referenced in congressional hearings, reports of the Congressional Research Service, internal and interagency memoranda in the federal executive, and other public records could easily overlook or understate its 'policy enlightenment' contributions. To the extent research might be used for an even broader 'public enlightenment' function, program evaluations might also require mechanisms and metrics for exploring the extent (if any) to which research findings gain footholds in public discourse. A challenging enough exercise at the time of the earliest studies of the agenda-setting function of the mass media,[31] tracking and tracing the diffusion of research language, findings, and ideas to public audiences is, with the proliferation of media channels, now considerably more complicated. Yet such efforts may be crucial in order to develop a full appreciation of the impacts of scholarly research. Consider funders' commitments to ensuring broad impacts of the research they support. Today best practice guidelines are widely available to assist investigators in: using Twitter in university research, teaching and impact activities; developing websites to share information about their work effectively; and employing social networking sites, blogging, video sharing, and content collecting and curating to communicate about their research (see, e.g., Mollett, Moran, and Dunleavy; Economic and Social Research Council, 2015). Efforts to explore the traction of research beyond the academy require evaluators to employ a similarly wide range of data and techniques—if use is an important criterion for valuing the research enterprise.

Also important from a practical perspective is that the many different uses that can be made of research evidence can take place over quite extensive periods of time (by multiple—targeted and unanticipated—audiences). This suggests data that could provide evidence of use need to be collected over considerable spans of time, if not longitudinally. 'Sleeping beauties'—scientific papers "whose relevance has not been recognized for decades, but then suddenly become highly influential and cited" (Ke, Ferrara, Radicchi, & Flammini, 2015: 7426) are a case in point. The issue is not a trivial one. Not only has recent examination of the sleeping beauty phenomenon suggested it is not as rare as one might think,[32] but even more proximal evidence of the use of scholarly findings is widely recognized, given the nature of publication and citation lag times, to be many years in the making. In some cases, the earliest projected 'uses' are not forecast to occur until after projects have concluded (or entire programs of investment have been sunsetted).

Another practical consideration is whether it is necessary, desirable, and/or feasible to weight use cases in some way. If, for example, the power an idea exerts when it results in an issue being identified as a priority for action on a stakeholder's agenda has consequences substantially greater than the influence it exerts when a single researcher or practitioner employs it in her work, it would seem this should be taken into account. It is not clear, however, how this would be accomplished—and what the implications would be for those charged with managing and evaluating portfolios of research investments. Should return on investment calculations weigh one 'powerful' instance more heavily than other instances of use? And what attempts should (and can) be made to distinguish 'good', 'appropriate' uses (e.g., based on sound understandings and defensible interpretations of study findings) from unsound, inappropriate ones—and account for these in weighting schemes?

Time is another important consideration in, and frequently an obstacle to, documenting use—particularly in the case of basic research studies. This is not the non sequitur it may at first seem. It has long been noted that research studies can achieve purposes far beyond those for which they were originally proposed. For example, late-stage effectiveness, scale-up, and other studies can be designed to document impact in broader contexts with larger populations than proof-of-concept or exploratory research, yet the former are as capable of generating the new ideas and contributions to theory that are typically associated with much earlier-stage foundational research. Similarly, individual 'basic' research studies, regardless of the investigators' intentions at the outset, may eventually be judged 'useful' and used—perhaps to generate new questions, perhaps to suggest new ideas for further research, perhaps in some other way—albeit often only when considered collectively and in the long-run. Indeed, as Bush argued in *Science: The Endless Frontier,*[33] while basic research may be conducted without consideration of practical ends, it is anticipated it will yield "general knowledge and an understanding of nature and its laws" which, in turn, "provides the means of answering a large number of important practical problems" (Bush, 1945: 18).

That such use occurs is fundamental to the notion that investments in basic research are socially desirable. As Nelson argues in his assessment of the economics of basic scientific research: "[f]rom a given expenditure on science we may expect a given flow, over time, of benefits that would not have been created had none of our resources been directed to basic research. This flow of benefits (properly discounted) may be defined as the social value of a given expenditure on basic research" (Nelson, 1959: 297). While the timeframe within which these 'practical ends' will be achieved is, at the outset, indeterminate, it is expected future uses will not be uncommon; "though many inventions are made possible by closely preceding advances in scientific knowledge, many others . . . occur long after the relevant scientific knowledge is available" (Nelson, 1959: 299). The important point is that use ultimately occurs, which could, in principle, be tied to some (not intolerably low) proportion of basic research endeavors.[34] The "scientist doing basic research may not be at all interested in . . . [its] practical applications" (Bush, 1945: 18). Nevertheless, her work—assuming it is accessible for consideration by others, not maintained solely in private papers that do not see the light of day—may well be used, in multiple ways, by multiple audiences, over time. Publicly available, non-proprietary, pure basic research holds the potential for use. Accordingly, such use is often taken into account in attempts to establish returns on foundational research investments.[35] Comprehensive efforts to link basic science to use are, however, currently beyond our reach. We have neither the tagging mechanisms nor the data infrastructure in place that would make it possible to trace and track distal impacts of research findings over time.

Challenges Inferring Value from Use

Implications for Evaluations of Investments in Education Research

Insights from decades of research on decision making and knowledge utilization have profound implications for evaluations of investments in education research. They draw attention to the numerous (design, measurement, and data-collection) challenges to documenting (and linking back to a specific research project or portfolio of investments) 'use'. They underscore the challenges of attributing any instances of use that might occur to the research investment (e.g., as opposed to investments in related yet separate initiatives to promote access to research findings). They also highlight the difficulties encountered in attempting to tease out the extent to which instances of use indicate something about the value attached to the research/research findings, versus something about the powerful influences other factors may have had on stakeholders' judgments. The challenges of inferring value from use are further complicated when similar or overlapping trajectories of research are supported by multiple funders and/or funding programs, and the goal of the evaluation is to comment on the value of or returns on specific programs of investment. These challenges are clearly illustrated when one attempts to sketch out a logic model that could guide a valuation with reference to use.

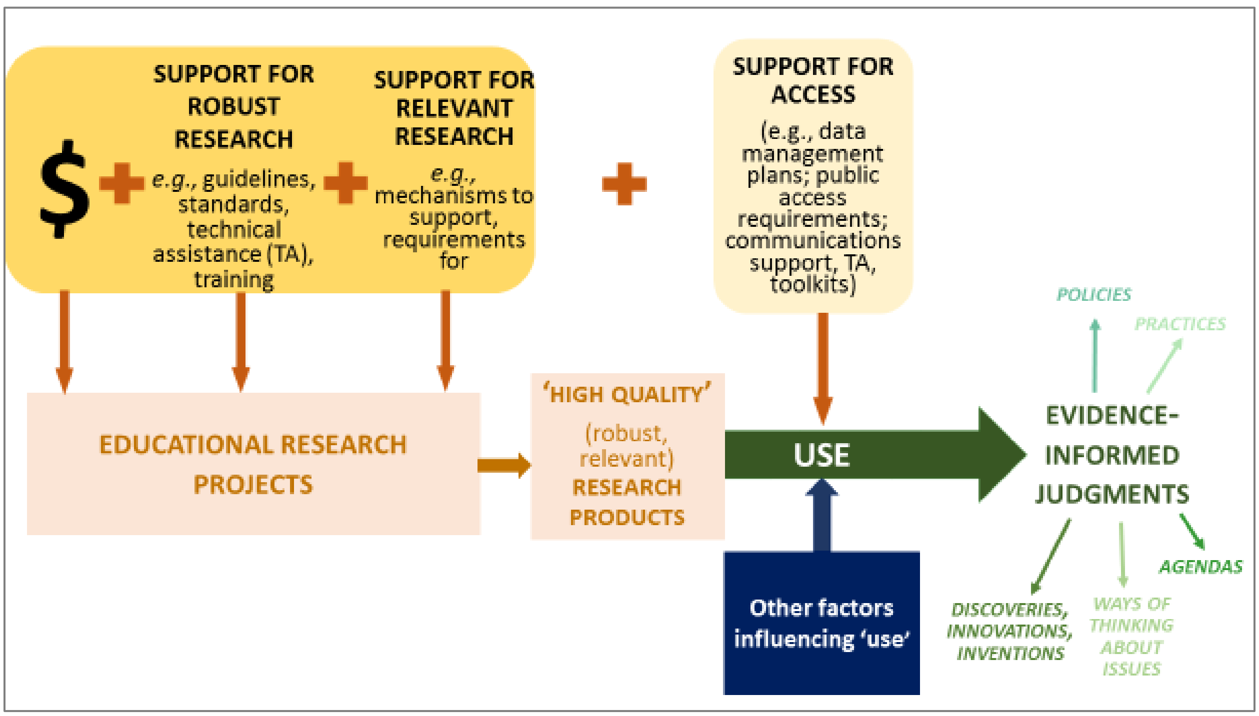
Logic models graphically portray the sequence of events through which it is hypothesized inputs contribute to sought impacts. They are visual representations of theories of action or change. The underlying theories of change map connections among project or program resources (e.g., funders' investments in education research), activities, outputs, outcomes, and impacts. In short, they clarify how it is expected change will occur.[36]

While logic models are frequently developed and employed for program management and evaluation purposes, an important, acknowledged, limitation of such models is their focus on a particular input-output chain; they typically portray a "single causal strand" of particular interest in light of the connection hypothesized between input and impact that drives program theory (see Renger, Wood, Williamson, and Krapp, 2011: 26). In principle, it is possible to draw a figure capturing all of the factors that might conceivably mediate or moderate inputs' ultimate influence on sought impacts; their inclusion can, however, considerably complicate the picture—potentially distracting attention from the theory of action the model portrays. Thus in practice these exogenous forces are often captured in an "external factors" or "context" bucket that acknowledges their potential importance to, but not the specific pathways by which they may enhance or disrupt, the inputs' effects.[37] Logic models and their underlying theories of action typically illuminate only key stages on the paths by which resources invested are expected to achieve sought outcome. Typically crafted to illustrate high-level features of a strategy or theory of action, they have the 'fast and frugal' qualities of heuristics (Gigerenzer & Goldstein, 1996: 650). Both are, by design, parsimonious; both (intentionally) fail to integrate all the information that could be helpful, focusing instead on capturing the most salient features for a particular decision-making purpose.

Useful as such simplified representations may be, it is important not to confuse them with actual representations of reality. Just as focusing on changes in particular metrics over time may place only a small subset of the (anticipated) outcomes of an education innovation under a magnifying glass, logic models may restrict focus detrimentally.[38] Moreover logic models are neither designed to illustrate counterfactual conditions, nor easily adapted to accommodate unanticipated changes in input-output trajectories (see NRC, 2014a: 53). Together, these characteristic features can at times prove challenging in designing and interpreting results of program evaluations.

The Exhibit below illustrates the challenges. Consistent with the priority federal policymakers have for nearly 20 years accorded efforts to enhance the rigor of and access to education research, it sketches out an emergent theory of action that distinguishes three 'inputs' influencing the conduct of research of: (1) funding dollars; (2) supports for (e.g., standards and guidelines for, and training, technical assistance, and other resources designed to build capacity to conduct) robust research; and (3) [CH1] s[MW2]upports for (e.g., requirements specifying prospective users should be consulted in prioritizing issues and identifying questions to be addressed in education research studies) relevant research. It also distinguishes the activities of research projects supported in this way from the products they produce. The disconnect between these products (the outputs of the funded research) and the impacts they may have (e.g., agendas; ways of thinking about issues; discoveries, innovations, and inventions; policies; practices) is graphically portrayed. The two are connected (and potentially separated) by 'use' behaviors, which, as illustrated, may be encouraged or discouraged by funder supports (efforts to support the accessibility of research products) and the wide range of additional factors the literature referenced above indicates shape judgment and decision making. The model highlights the roles these supports and factors play in determining which research findings are ultimately 'used'. It suggests that the research evidence's latent potential to inform the judgments must be 'activated' to yield the (more or less timely, proximal and distal, direct and/or indirect) instances of use that are prized when use is employed as a criterion for valuing education research. No effort is made, however, to enumerate or illustrate the multiple paths by which these factors may shape the decisions that lead to these impacts.

Exhibit. An emergent theory of action for transforming education research investments into valued outcomes



Recreating Expectations

Roles for Program Evaluators

Funders are understandably eager to ensure the moneys with which they have been entrusted are well spent. To this end they offer guidance to applicants and awardees—signaling, when not prescribing, the objectives they intend their investments to realize. The extent to which funded projects, individually or collectively, are successful in achieving these objectives (i.e., have the desired impacts) would thus seem a reasonable, if not the only, basis on which to gauge the wisdom and value of their investments. It would also seem helpful in warranting continuation of a particular investment strategy or approach. Yet as the simple, generic, skeletal logic model sketched above illustrates, funders' inputs may be only indirectly connected to their desired outcomes (e.g., research being used to inform education reform) and impacts (e.g., in the short term, the adoption of a curriculum with demonstrated effectiveness when used with a particular population of learners; in the long run, increases in those students' learning outcomes). Other factors may come into play (e.g., district administrators' approaches to problem framing, experience, local knowledge, political considerations, personal beliefs and values), mitigating or moderating the extent to which findings of high-quality research influence their judgments. Rigorous, relevant research clearly has the potential to inform reform. When steps designed to encourage research utilization are of particular interest, it is appropriate to find them highlighted in program logic models. At the same time it is important not to lose sight of the many other factors likely to influence the outcomes of interest; doing so risks drawing inappropriate conclusions regarding the return on research investments.

Evaluators have important roles to play in ensuring presumptions regarding the use of research findings do not result in inappropriate conclusions being drawn regarding the value of and returns on the research enterprise. First, they can highlight and clarify the complex, dynamic, and often decider-specific processes through which rigorous, relevant, accessible research findings might be employed—by whom, and within what timeframes. As extensive—and intuitively appealing—as extant research findings on these issues are, they are apparently not well appreciated by some who position education research as ***"the"*** driver of education reform. Clearly, research evidence has the potential to encourage 'sound', informed education policies and practices. Whether or not this potential will be activated may depend less on the soundness or the quality of the research enterprise, more on factors far beyond either the funders' or the investigators' control. From individual scholar to funder-public information officer, many individuals outside the domain of specific funding programs have key roles to play in positioning research products as potential inputs to others' 'informed' decisions. To the extent the steps necessary to move research products from output to impact by way of use are not within the jurisdiction of individual programs or their managers, it would be inappropriate to downgrade assessments of the quality, utility, or value of research based on an apparent absence of use. If use is assumed indicative of impact(s) of interest, efforts to press evidence higher up the list of priority considerations influencing education administrators' and educators' decisions may be key mediators, meriting fuller consideration.

Evaluators can help establish the prospects—and consider the costs—of tracking use. As alluded to above, numerous methodological, measurement, and data challenges are encountered in efforts to document use of research evidence—and to ascribe use appropriately to individual studies or portfolios of investment. Issues of timing are critically important here. For example, if tracing research to the broadest possible impacts is their goal, program managers may very well need to add to their statements of work (and budgets) for evaluation contracts requirements that offerors have experience employing social media and web analytics to monitor the reach and effects of digital communications. Given the long timeframes within which various instances of use are acknowledged to take place, there are dangers in placing premature closure on evaluative studies. At the same time, the costs of monitoring key deciders' actions and records of their decisions can be prohibitive. Costs can mount even higher when the dynamic nature of programs (e.g., shifting priorities) suggests evidence of use should be sought among additional or other target audiences over time, or when they require post hoc reexaminations of longitudinal data.

Evaluators can remind those who commission their work of the limitations of overly elegant logic models. There are numerous benefits to generating simplified models of theories of change; they focus attention on factors assumed, hypothesized, or presumed to be sufficiently important that they are taken into account in developing and directing programs of investment. This very simplicity is also one of their acknowledged limitations. Attempting to gauge return on investment based solely on the factors highlighted in the simple model sketched out above, for example, assumes a comprehensiveness the model (by design) lacks. While not all statements of work are amenable to the suggestion, it may be wise for evaluators to press for resources to conduct an evaluability assessment collaboratively with the funder before a scope of work is finalized. This would ensure both readiness for evaluation and that intended purposes and anticipated uses of the results of the evaluation are feasible—from design and cost perspectives.[39]

Evaluators can also remind funders (and those who judge the wisdom of their funding decisions) of the merits of output evaluations. Given the number of factors beyond funders' control which may enhance, impede, or constrain use, it may be preferable in warranting investment strategies to focus specifically on the steps toward impact they can reasonably be expected to leverage (thus be held accountable for).

Closer consideration of nascent logic models may suggest it is most appropriate to gauge the contributions, value, and returns on investments in education research with an eye to the outputs (e.g., research products) studies yield. These are arguably most amenable to funders' influence (e.g., through the direction and guidance they provide in their notices of funding opportunities and their other interactions with the field; the processes they employ to establish the merits of proposed investigations and recommend award decisions; the attention they devote to post-award management responsibilities). Pragmatically, because they are proximal to funding decisions and more closely tied to factors within the funders' sphere of influence, output-focused evaluations are often both more informative and motivating for program managers. There may be little a program manager can (or should) do to influence the ultimate impacts (use or otherwise) associated with education research funding decisions; there may be much she can do, however, to ensure promised deliverables are provided according to project timelines, and agreements entered into regarding steps that would be taken to assure the rigor, relevance, and accessibility of research products are adhered to. If the means-end chain sketched out in the program's logic model holds, these may well not be the final objectives of a portfolio of investments—but they should be critical to their achievement.

Finally, evaluators can, and perhaps should, encourage funders and science policymakers alike to reconsider the benefits and drawbacks—for investigators who conduct, program staff who manage, and organizations that support education research—of employing use as a criterion for assessing its value. The culture of evidence and accountability are such that federal agencies will continue to seek evidence of impact (or something as close to it as can be mustered). Yet the challenges of conducting impact evaluations—whether from a 'utilization' or other perspective—are well understood not only by evaluators but by those who commission their work. The challenge is to provide meaningful benchmarks and measures suggesting investments of public funds are prudently and appropriately directed and managed. U.S. millennial education reforms have embraced and promulgated a "research can fix it" culture. Appropriately channeled, the value such a culture places on the research enterprise may stimulate potentially transformative innovations, discoveries, and learning environments. However, left unchecked, it may also foster unrealistic and potentially damaging expectations regarding the roles education research can play in education reform, and the appropriateness of employing evidence of use as a criterion for assessing the impacts of and the returns on federal investments in education research. Expectations that high-quality research could and should be employed to resolve persistent challenges improving student learning and the equality of educational opportunity may have fostered the impression that they would. Sustained support for careful attention to the factors influencing quality and accessibility of education research may well have had substantial, meaningful impacts on both, but there is no guarantee that even the most rigorous, relevant, and accessible research findings will be used to a particular effect.

The above is not meant to imply that federal policymakers' NCLB era efforts to enhance the quality of and access to robust research findings were in any way misguided. Rigorous, relevant, easily accessible study findings may be more likely to affect policymakers' and practitioners' decisions and courses of action—but decades of research and practical experience indicate the rigor, relevance, and accessibility of research findings are neither necessary nor sufficient to ensure research products will be used. Interesting or novel findings of studies with limited or poor scientific value can still capture the imagination and find their way into the classroom (e.g., informing instructional practice) or a school improvement plan. The factors influencing the extent (if any) to which research findings are utilized to inform decision-makers' and practitioners' beliefs, values, thoughts, and actions are many, varied, and complex. Moreover, many are out of the control, or even the sphere of influence, of those who fund and conduct the research.

Over 50 years have elapsed since the Center for Research on the Utilization of Scientific Knowledge was established at the University of Michigan's Institute for Social Research in 1964 to "address the failure of organizations and individuals to make use of scientific knowledge" (Institute for Social Research, 2015),[40] nearly 40 since an NRC report on *Knowledge and Policy: The Uncertain Connection* found systematic evidence on whether efforts to enhance the utility of research were "having the results their sponsors hope for" lacking (NRC, 1978: 5). Today research on research utilization continues to be a focus of scholarly endeavor and public interest. Progress has been made, but efforts to advance understanding of the ways in which research is used to "enrich problem framing, decision-making, and individual and organizational learning in education" (Tseng & Nutley, 2014: 173) continue to be characterized as "still a relatively young field of study" (Tseng, 2012: 12). Policymakers and scholars continue to be exhorted to undertake investigations that will foster better understanding of the conditions supporting the use of social science knowledge in policymaking (see e.g., NRC, 2012). At the same time, state and local administrators and classroom teachers continue to be expected to develop and follow both evidence-based strategies for improving the induction, mentoring, professional development, and retention of educators; and evidence-based strategies and interventions for improving student achievement, instruction, and schools (see e.g., the *Every Student Succeeds Act*, Pub.L. 114-95). It is gratifying to see the potential role education research can play in informing reform given such recognition. It would be unfortunate if failure to overcome the many obstacles to ensuring evidence informs judgment fueled questions regarding the merits of continued financial, and other, supports for the research enterprise. Failure to find evidence of or to trace the uses made of high-quality, accessible research findings should not be assumed to be an indictment of the underlying research.

References

Averch, Harvey A.; Carroll, Stephen J.; Donaldson, Theodore S.; Kiesling, Herbert J.; & Pincus, John. (1972). *How Effective Is Schooling: A Critical Review and Synthesis of Research Findings*. Prepared for the President's Commission on School Finance. R-956-PCSF/RC. Santa Monica, CA: The Rand Corporation. Retrieved September 2016 from http://www.rand.org/content/dam/rand/pubs/reports/2006/R956.pdf.

Bazerman, Max; & Moore, Don A. (2012). *Judgment in Managerial Decision Making*, 8th edition. New York: John Wiley & Sons.

Beyea, Suzanne C.; & Slattery, Mary Jo. (2013). Historical perspectives on evidence-based nursing. *Nursing Science Quarterly, 26*(2): 152-155.

Beyer, Janice M. (1997). Research utilization: Bridging a cultural gap between communities. *Journal of Management Inquiry, 6*(1): 17-22.

Bryk, Anthony S. (2009). Supporting a science of performance improvement. *Phi Delta Kappan, 90*(8): 597-600.

Bryk, Anthony S. & Gomez, Louis M. (2008). Ruminations on reinventing an R&D capacity for educational improvement. In F.M. Hess (Editor) *The Future of Educational Entrepreneurship: Possibilities of School Reform*. Cambridge, MA: Harvard Education Press.

Bryk, Anthony S.; Gomez, Louis M., & Grunow, Alicia. (2010). Getting ideas into action: Building networked improvement communities in education. *Carnegie Perspectives*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved September 2016 from http://cdn.carnegiefoundation.org/wp-content/uploads/2014/09/bryk-gomez_building-nics-education.pdf.

Burwell, Sylvia M.; Munoz, Cecilia; Holdren, John; & Krueger, Alan. (2013). Next steps in the evidence and innovation agenda. Memorandum to the Heads of Executive Departments and Agencies, M-13-17. Washington, D.C.: Executive Office of the President, Office of Management and Budget. Retrieved September 2016 from https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-17.pdf.

Bush, V. (1945). *Science: The Endless Frontier*, Report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development. Washington, DC: United States Government Printing Office. Retrieved September 2016 from https://www.nsf.gov/od/lpa/nsf50/vbush1945.htm.

Caplan, N. (1979). The two-communities theory and knowledge utilization. *American behavioral Scientist 22*(3): 459-470.

Claridge, J.A., and Fabian, T.C. (2005). History and development of evidence-based medicine. *World Journal of Surgery, 29*(5): 547-553.

Cobb, R.W., and Elder, Y.C. (1983). Participation in American politics: The dynamics of agenda-building, 2nd edition. Baltimore: The Johns Hopkins University Press.

Cobb, R.W., and Ross, M.W. (1997). Agenda setting and the denial of agenda access: Key concepts. In R.W. Cobb & M.H. Ross (Eds.), *Cultural Strategies of Agenda Denial*. Lawrence: University Press of Kansas.

Cobb, R.W., Ross, J-K., and Ross, M.H. (1976). Agenda building as a comparative political process. *American Political Science Review, 70*(1): 126-138.

Coburn, Cynthia E.; Honig, M.I.; & Stein, M.K. (2009). What's the evidence on districts' use of evidence? In J. Bransford, D.J. Stipek, N.J. Vye, L. Gomez, and D. Lam (Eds.) *The Role of Research in Educational Improvement.* Cambridge, MA: Harvard Education Press.

Cochrane, Archibald L. (1972). *Effectiveness and Efficiency: Random Reflections on Health Services*. Oxford, England: The Nuffield Provincial Hospitals Trust. Retrieved September 2016 from http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/Effectiveness_and_Efficiency.pdf.

Committee on Science, Engineering, and Public Policy (COSEPUP). (1986). Report of the Research Briefing Panel on Decision Making and Problem Solving. In *Research Briefings 1986*. Washington, D.C.: National Academy Press.

Corcoran, Thomas B. (2003). The use of research evidence in instructional improvement. *CPRE Policy Briefs*, RB-40. Philadelphia, PA: Consortium for Policy Research in Education (CPRE). Retrieved September 2016 from http://repository.upenn.edu/cpre_policybriefs/27.

Cousins, J.B., & Leithwood, K.A. (1993). Enhancing knowledge utilization as a strategy for school improvement. *Knowledge: Creation, Diffusion, Utilization, 14*(3): 305-333.

Davies, Huw T.O., and Nutley, Sandra M. (2008). Learning More About How Research-Based Knowledge Gets Used: Guidance in the Development of New Empirical Research, Working Paper. New York, NY: William T. Grant Foundation. Retrieved September 2016 from http://www.ruru.ac.uk/pdf/WT%20Grant%20paper_final.pdf.

Davis, H., Nutley, S., & Walter, I. (2005). *Approaches to assessing the non-academic impact of social science research*, Report of an ESRC [Economic and Social Research Council] Symposium May 12-13, 2005. Swindon, England: Economic and Social Research Council. Retrieved September 2016 from http://www.esrc.ac.uk/files/research/evaluation-and-impact/international-symposium/.

Donovan, Shaun; & Holdren, John. (2015). Multi-agency science and technology priorities for the F 2017 budget, Memorandum for the heads of executive departments and agencies (M-15-16). Washington, D.C.: Executive Office of the President, Office of Management and Budget. Retrieved September 2016 from https://www.whitehouse.gov/sites/default/files/omb/memoranda/2015/m-15-16.pdf.

Downs, A. (1972). Up and Down with Ecology: The Issue-Attention Cycle. *Public Interest, 28*: 38-50.

Economic and Social Research Council (ESRC). (2015). Impact Toolkit. Swindon, England: Economic and Social Research Council. Retrieved September 2016 from http://www.esrc.ac.uk/research/impact-toolkit/.

Education Sciences Reform Act of 2002, Pub. L. No. 107-279 (2002). Retrieved September 2016 from http://www2.ed.gov/policy/rschstat/leg/PL107-279.pdf

Eidell, T. L. & Kitchel, J. M. (Eds.) (1968). *Knowledge Production and Utilization in Educational Administration*. Eugene, OR: Center for Advanced Study of Educational Administration, University of Oregon.

Estabrooks, Carole A.; Derksen, Linda; Winter, Connie; Lavis, John N.; Scott, Shannon D.; Wallin, Lars; and Profetto-McGrath, Joanne. (2008). The intellectual structure and substance of the knowledge utilization field: A longitudinal author co-citation analysis, 1945 to 2004. *Implementation Science, 3*: 49-70.

Fealing, Kaye Husbands; Lane, Julia I.; Marburger, John H. III; Shipp, Stephanie S. (2011). *The Science of Science Policy: A Handbook*. Stanford, CA: Stanford University Press.

Feuer, Michael J.; Towne, Lisa; and Shavelson, Richard J. (2002). Scientific culture and educational research. *Educational Researcher, 31*(8): 4-14.

Finnigan, Kara S., & Daly, Alan J. (Eds.). (2014). *Using Research Evidence in Education: From the Schoolhouse Door to Capitol Hill*. Heidelberg: Springer International Publishing.

Fowler, P.B.S. (1997). Evidence-based everything. *Journal of Evaluation in Clinical Practice, 3*(3): 239-243.

Frechtling, J.A. (2007). *Logic Modeling Methods in Program Evaluation*. San Francisco, CA: Jossey-Bass.

Funnell, Sue C., & Rogers, Patricia J. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco, CA: Jossey-Bass.

Gigerenzer, Gerd; & Goldstein, Daniel G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4): 650-669.

Gino, Francesca. (2013). *Sidetracked: Why Our Decisions Get Derailed, and How We Can Stick to the Plan*. Cambridge, MA: Harvard Business Review Press.

Gino, Francesca; Brooks, Alison Wood; and Schweitzer, Maurice E. (2012). Anxiety, advice, and the ability to discern: Feeling anxious motivates individuals to seek and use advice. *Journal of Personality and Social Psychology, 102*(3): 497-512.

Gladwell, Malcolm. (2005). *Blink: The Power of Thinking Without Thinking*. Boston, MA: Little, Brown & Company.

Hood, Paul. (2002). Perspectives on knowledge utilization in education. San Francisco, CA: WestEd. Retrieved September 2016 from https://www.wested.org/resources/perspectives-on-knowledge-utilization-in-education/

Hummelbrunner, Richard. (2011). Beyond logframe: Critique, variations and alternatives. In Nobuko Fujita (Ed.) *Beyond Logframe: Using Systems Concepts in Evaluation*, Issues and Prospects of Evaluations for International Development, Series IV. Tokyo, Japan: Foundation for Advanced Studies on International Development. Retrieved September 2016 from http://www.seachangecop.org/sites/default/files/documents/2010%2003%20FASID%20-%20Beyond%20Logframe%20-%20Using%20systems%20concepts%20in%20evaluation.pdf.

Institute of Education Sciences (IES), U.S. Department of Education. (2008). *Rigor and Relevance Redux: Director's Biennial Report to Congress*, IES 2009-6010. Washington, D.C.: U.S. Department of Education. Retrieved September 2016 from http://ies.ed.gov/director/pdf/20096010.pdf.

Institute for Social Research. (2015). Institute for Social Research: Timeline. Ann Arbor, MI: Institute for Social Research, University of Michigan. Retrieved September 2016 from http://home.isr.umich.edu/about/history/timeline/.

Jones, Molly Morgan; Castle-Clarke, Sophie; Manville, Catriona; Gunashekar, Salil; and Grant, Jonathan. (2013). Assessing research impact: An international review of the Excellence in Innovation for Australia Trial. Cambridge, UK: RAND Europe. Retrieved September 2016 from http://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR278/RAND_RR278.pdf.

Kahneman, Daniel. (2011). *Thinking, Fast and Slow*. New York: Farrar, Strauss, Giroux.

Kahneman, Daniel. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Kahneman, Daniel & Tversky, Amos, (Eds.). (2000). *Choices, Values and Frames*. New York: Cambridge University Press and the Russell Sage Foundation.

Kahneman, Daniel; Slovic, P., & Tversky, Amos. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Ke, Qing; Ferrara, Emilio; Radicchi, Filippo; & Flammini, Alessandro. (2015). Defining and identifying Sleeping Beauties in science. *Proceedings of the National Academy of Sciences, 112*(24): 7426-7431.

Kingdon, J.W. (1995). *Agenda, Alternatives and Public Policies*. New York: HarperCollins Publishers.

Kirst, Michael W. (2000). Bridging education research and education policymaking. *Oxford Review of Education*, *26*(3/4): 379-391.

Knott, J., and Wildavsky, A. (1980). If dissemination is the solution, what is the problem? *Knowledge: Creation, Diffusion, Utilization, 1*(4): 537-578.

Kohlmoos, Jim & Joftus, Scott. (2005). Building communities of knowledge: Ideas for more effective use of knowledge in education reform. Discussion paper. Washington, D.C.: NEKIA.

Langdon, David; McKittrick, George; Beede, David; Khan, Beethika; & Doms, Mark. (2011). *STEM: Good Jobs Now and for the* Future, ESA [Economics and Statistics Administration] Issue Brief #03-11. Washington, D.C.: U.S. Department of Commerce, Economics and Statistics Administration. Retrieved September 2016 from http://files.eric.ed.gov/fulltext/ED522129.pdf.

Larsen, Judith K. (1980). Knowledge utilization: What is it? *Science Communication*, *1*(3): 421-442.

Lehming, Rolf & Kane, Michael B. (1981). *Improving Schools: Using What We Know*. Beverly Hills, CA: Sage Publications.

Levin, Ben. (2013). What does it take to scale up innovations? An examination of Teach for America, the Harlem Children's Zone, and the Knowledge is Power Program. Policy Brief, National Education Policy Center. Boulder, CO: National Education Policy Center, School of Education, University of Colorado. Retrieved September 2016 from http://nepc.colorado.edu/publication/scaling-up-innovations.

Lindblom, Charles E., & Cohen, David K. (1979). *Usable Knowledge: Social Science and Social Problem Solving*. New Haven, CT: Yale University Press.

LSE Public Policy Group. (2008). *Maximizing the Social, Policy and Economic Impacts of Research in the Humanities and Social Sciences: Research Report*. Supplementary Report to the British Academy from the LSE Public Policy Group. London, England: London School of Economics and Political Science.

Lynd, R.S. (1939). *Knowledge for What? The Place of Social Science in American Culture*. Princeton, NJ: Princeton University Press.

Marburger, J.H., III. (2005a). Speech made April 21, 2005 at the 30th Annual AAAS Forum on Science and Technology Policy in Washington, D.C. Retrieved September 2016 from http://www.scienceofsciencepolicy.net/sites/default/files/attachments/2005-04-21%20Speech%20-%20to%20AAAS%20Forum%20on%20Science%20and%20Technology%20Policy_0.pdf.

Marburger, J.H., III. (2005b). Wanted: Better benchmarks. *Science*, *308*(5725): 1087.

McCombs, Maxwell E. & Shaw, Donald. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, *36*(2): 176-187.

McCombs, Maxwell. (2004). *Setting the Agenda: The Mass Media and Public Opinion*. Cambridge, England: Polity Press.

McDonald, S-K. (2009). Scale-up as a framework for program and policy evaluation research. In David N. Plank, Barbara Schneider, and Gary Sykes (Eds.). *Handbook on Education Policy Research*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

McDonald, S-K., Keesler, V., Kauffman, N., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher 35*(3):15-24.

McLaughlin, J.A., & Jordan, G.B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Planning*, *22*(1): 65-72.

Meagher, L.R. (2013). Research Impact on Practice: Case Study Analysis. Swindon, England: Economic and Social Research Council. Retrieved February 2015 from http://www.esrc.ac.uk/_images/Research-impact-on-practice_tcm8-25587.pdf.

Members of the 2005 "Rising Above the Gathering Storm" Committee. (2010). *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5*. Prepared for the Presidents of the National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. Washington, D.C.: The National Academies Press.

Millar, Annie; Simeone, Ronald S.; & Carnevale, John T. (2001). Logic models: A systems tool for performance management. *Evaluation and Program Planning*, *24*(1): 73-81.

Mollett, Amy; Moran, Danielle; & Dunleavy, Patrick. (2011). Using Twitter in university research, teaching and impact activities: A guide for academics and researchers. London, England: LSE Public Policy Group. Retrieved September 2016 from http://blogs.lse.ac.uk/impactofsocialsciences/files/2011/11/Published-Twitter_Guide_Sept_2011.pdf.

National Education Knowledge Industry Association (NEKIA). (2005). NEKIA unveils new principles for knowledge use in school improvement: Principles support the launch of new era of improvement &

innovation in education. Press release dated June 28, 2005. Washington, D.C.: NEKIA. Retrieved September 2016 from http://www.knowledgeall.net/files/knowledge_principles062805.pdf.

National Research Council (NRC). (2014a). *Furthering America's Research Enterprise*. R.F. Celeste, A. Griswold, and M.L. Straf (Editors). Committee on Assessing the Value of Research in Advancing National Goals, Division of Behavioral and Social Sciences and Education. Washington, D.C.: The National Academies Press.

National Research Council (NRC). (1978). *Knowledge and Policy: The Uncertain Connection*. Laurence E. Lynn, Jr. (Ed.). Study Project on Social Research and Development, Volume 5. Washington, D.C.: National Academy of Sciences, Printing and Publishing Office.

National Research Council (NRC). (1992) *Research and Education Reform: Roles for the Office of Educational Research and Improvement*. Committee on the Federal Role in Education Research. Richard C. Atkinson and Gregg B. Jackson (Eds.). Washington, D.C.: The National Academies Press.

National Research Council (NRC). (2007). *Rising Above the Gathering Storm: Energizing and Employing American for a Brighter Economic Future*. Committee on Prospering in the Global Economy of the 21st Century and the Committee on Science, Engineering, and Public Policy. Washington, D.C.: The National Academies Press.

National Research Council (NRC). (2002). *Scientific Research in Education*. Committee on Scientific Principles for Education Research. Shavelson, R.J., and Towne, L., (Eds.). Center for Education. Division of Behavioral and Social Sciences and Education. Washington, D.C.: The National Academies Press.

National Research Council (NRC). (2012). *Using Science as Evidence in Public Policy*. Committee on the Use of Social Science Knowledge in Public Policy. K. Prewitt, T.A. Schwandt, and M.L. Straf, Editors. Division of Behavioral and Social Sciences and Education. Washington, D.C.: The National Academies Press.

National Science Board. (2015). *Revisiting the STEM Workforce: A companion to Science and Engineering Indicators 2014*, NSB-2015-10. Arlington, VA: National Science Foundation. Retrieved September 2016 from http://www.nsf.gov/nsb/publications/2015/nsb201510.pdf.

National Science Foundation (NSF). (2015). Discovery Research PreK-12 (DRK-12) Program Solicitation (NSF 15-592). Arlington, VA: National Science Foundation. Retrieved September 2016 from http://www.nsf.gov/pubs/2015/nsf15592/nsf15592.pdf.

National Science Foundation. (2006). *National Science Foundation FY 2007 Budget Request to Congress*. Arlington, VA: National Science Foundation. Retrieved September 2016 from http://www.nsf.gov/about/budget/fy2007/.

National Science Foundation (NSF) and U.S. Department of Education, Institute of Education Sciences (IES). (2013). *Common Guidelines for Education Research and Development*, NSF 13-126. Arlington, VA: National Science Foundation. Retrieved September 2016 from http://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf.

Neilson, Stephanie. (2001). IDRC-supported research and its influence on public policy: Knowledge utilization and public policy processes—a literature review. Ottawa, Canada: Evaluation Unit, International Development Research Centre (IDRC).

Nelson, Richard R. (1959). The simple economics of basic scientific research. *Journal of Political Economy*, *67*(3): 297-306. Retrieved September 2016 from http://cstpr.colorado.edu/students/envs_5100/nelson_1959.pdf.

Newell, A., and Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

No Child Left Behind Act of 2001, Pub. L. No. 107-111, § 115, Stat. 1425 (2002). Retrieved September 2016 from http://www.gpo.gov/fdsys/pkg/PLAW-107publ110/pdf/PLAW-107publ110.pdf.

Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge, England: Polity Press.

Nutley, S.M., Walter, I., and Davies, H.T.O. (2007). *Using Evidence: How Research Can Inform Public Services*. Bristol, England: The Policy Press.

Oakley, Ann. (2002). Social science and evidence-based everything: The case of education. *Educational Review*, *54*(3): 277-286. Retrieved September 2016 from http://www.cebma.info/wp-content/uploads/Oakley-Social-Science-and-Evidence-based-Everything-The-case-of-education.pdf.

Orszag, P.R., & Holdren, J.P. (2009). Science and Technology Priorities for the FY 2011 Budget: Memorandum for the heads of executive departments and agencies. M-09-27. Washington, D.C.: The White House. Retrieved September 2016 from http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_fy2009/m09-27.pdf

Paisley, W. J. & Butler, M. (1983). *Knowledge Utilization Systems in Education: Dissemination, Technical Assistance, Networking*. Beverly Hills, CA: Sage Publications.

Pelz, D.C. (1978). Some expanded perspectives on the use of social science in public policy. In J.M. Yinger and S.J. Cutler (Editors) *Major Social Issues: A Multidisciplinary View*. New York, NY: Free Press.

Peters, B.G. (2007). Agenda setting and public policy. In B.G. Peters (Ed.), *American Public Policy: Promise and Performance*. Washington, D.C.: CQ Press.

Pettigrew, A.M. (2011). Scholarship with impact. *British Journal of Management*, *22*(3): 347-354.

RAND Europe. (2012). Impact and the Research Excellence Framework: New challenges for universities. Santa Monica, CA: RAND Corporation. Retrieved September 2016 from http://www.rand.org/content/dam/rand/pubs/corporate_pubs/2012/RAND_CP661.pdf.

Renger, Ralph; Wood, Shandiin; Williamson, Simon; Krapp, Stefanie. (2011). Systematic evaluation, impact evaluation and logic models. *Evaluation Journal of Australasia*, *11*(2): 24-30.

Rich, Robert F. (1997). Measuring knowledge utilization: Processes and outcomes. *Knowledge and Policy*, *10*(3): 11-24.

Roderick, Melissa; Easton, John Q.; & Sebring, Penny Bender. (2009). *The Consortium on Chicago School Research: A New Model for the Role of Research in Supporting Urban School Reform*. Chicago, IL: Consortium on Chicago School Research at the University of Chicago Urban Education Institute. Retrieved September 2016 from http://ccsr.uchicago.edu/sites/default/files/publications/CCSR%20Model%20Report-final.pdf.

Russo, J.E., and Schoemaker, P.J.H. (1989). *Decision Traps: The Ten Barriers to Brilliant Decision-Making and How to Overcome Them*. New York, NY: Simon & Schuster.

Schattschneider, E.E. (1960). *The Semisovereign People*. New York: Holt, Rinehart and Winston.

Scientifically Based Education Research, Statistics, Evaluation, and Information Act of 2000. H.R. 4875, 106th Congress. (2000). Retrieved September 2016 from https://www.congress.gov/106/bills/hr4875/BILLS-106hr4875ih.pdf

Simon, Herbert A. (1972). Theories of bounded rationality. In C.B. McGuire and R. Radner (Eds.), *Decision and Organization* (pp. 161-176). Amsterdam, The Netherlands: North-Holland.

Simon, Herbert A. (1960). *The New Science of Management Decision*. New York, NY: Harper & Brothers Publishers.

Smith, M.F. (1989). *Evaluability Assessment: A Practical Approach*. New York, NY: Springer Science + Business Media.

Soydan, Haluk; and Palinkas, Lawrence A. (2014). *Evidence-based Practice in Social Work: Development of a New Professional Culture*. New York, NY: Routledge.

Starkey, K., and Madan, P. (2001). Bridging the relevance gaps: Aligning stakeholders in the future of management research. *British Journal of Management, 12*(s1): S3-S26.

Tallman, I., Leik, R.K., Gray, L.N., and Stafford, M.C. (1993). A theory of problem-solving behavior. *Social Psychology Quarterly, 56*(3): 157-177.

Taylor-Powell, E., Jones, L., & Henert, E. (2003). *Enhancing Program Performance with Logic Models*. Retrieved September 2016 from the University of Wisconsin-Extension web site: http://www.uwex.edu/ces/lmcourse/.

Tetlock, P.E. (2005). *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton: Princeton University Press.

Trevisan, Michael S. (2007). Evaluability assessment from 1986 to 2006. *American Journal of Evaluation, 28*(3): 290-303.

Trevisan, Michael S.; & Huang, Yi Min. (2003). Evaluability assessment: A primer. *Practical Assessment, Research & Evaluation, 8*(20). Retrieved January 2016 from http://pareonline.net/getvn.asp?v=8&n=20.

Tseng, Vivian. (2012). The uses of research in policy and practice. *Social Policy Report, 26*(2): 1-16.

Tseng, Vivian; & Nutley, Sandra. (2014). Building the infrastructure to improve the use and usefulness of research in education. In K.S. Finnigan and A.J. Daly (Eds.) *Using Research Evidence in Education: From the Schoolhouse Door to Capitol Hill*, Policy Implications of Research in Education 2. Switzerland: Springer International Publishing.

U.S. Department of Education. (2015a). Funding opportunities: Education research grant programs [online]. Washington, D.C.: U.S. Department of Education. Retrieved September 2016 from http://ies.ed.gov/funding/ncer_rfas/postdoc_training.asp.

U.S. Department of Education. (2015b). Request for Applications, Education Research Grants, CFDA Number 84.305A. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved September 2016 from http://ies.ed.gov/funding/pdf/2016_84305A.pdf.

U.S. Department of Education. (2014). Request for Applications, Education Research and Development Center Program, CFDA Number 84.305C. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved September 2016 from http://ies.ed.gov/funding/pdf/2015_84305C.pdf.

U.S. Department of Education. (2013). *Institute of Education Sciences Director's Biennial Report to Congress, Fiscal Years 2011 and 2012*, IES 2014-001. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved September 2016 from http://ies.ed.gov/pdf/2014001.pdf.

U.S. Department of Education, Office of the Under Secretary. (2002). *No Child Left Behind: A Desktop Reference, 2002*. Washington, D.C.: U.S. Department of Education. Retrieved September 2016 from https://www2.ed.gov/admins/lead/account/nclbreference/reference.pdf.

U.S. Department of Justice, National Institute of Corrections. (2013). *Evidence-Based Practices in the Criminal Justice System: Annotated Bibliography*. Washington, D.C.: U.S. Department of Justice. Retrieved September 2016 from http://static.nicic.gov/Library/026917.pdf.

U.S. House of Representatives. Committee on Education and the Workforce. (2002). President Bush signs landmark education reforms into law. Press release, January 8, 2002. Washington, D.C.: U.S. Congress, House of Representatives. Retrieved September 2016 from http://archives.republicans.edlabor.house.gov/archive/press/press107/hr1signing10802.htm.

U.S. President's Commission on School Finance. (1971). *Progress Report of the President's Commission on School Finance*. Washington, D.C.: President's Commission on School Finance. Retrieved September 2016 from http://files.eric.ed.gov/fulltext/ED058643.pdf.

U.S. President's Commission on School Finance. (1972). *Schools, People, & Money: The Need for Educational Reform*. Final Report, President's Commission on School Finance, Neil Hosler McElroy. Washington, D.C.: U.S. Government Printing Office.

W.K. Kellogg Foundation. (2004). *Using Logic Models to Bring Together Planning, Evaluation, and Action: Logic Model Development Guide*. Battle Creek, MI: W.K. Kellogg Foundation. Retrieved September 2016 from http://www.smartgivers.org/uploads/logicmodelguidepdf.pdf.

Weiss, Carol H. (1977). Research for policy's sake: the enlightenment function of social research. *Policy Analysis, 3*(4): 531-545.

Weiss, Carol H. (1979). The many meanings of research utilization. *Public Administration Review, 39*(5): 426-431.

Weiss, C.H. (1988). Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice, 9*(1): 5-19.

Weiss, C.H. (1995). The Haphazard connection: Social science and public policy. *International Journal of Educational Research, 23*(2): 137-150.

White House. (2002). Fact Sheet: No Child Left Behind Act. Washington, D.C.: White House, Executive Office of the President. Retrieved September 2016 from http://georgewbush-whitehouse.archives.gov/news/releases/2002/01/print/20020108.html.

Whitehurst, Grover J. ("Russ"). (2008). *Rigor and Relevance Redux,* Director's Biennial Report to Congress, IES 2009-6010. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved September 2016 from http://files.eric.ed.gov/fulltext/ED503384.pdf.

Wholey, Joseph S. (2004). Evaluability assessment: Developing program theory. *New Directions for Evaluation, 1987*(3): 77-92.

William T. Grant Foundation. (2015). Focus areas: Improving the use of research evidence. New York, NY: William T. Grant Foundation. Retrieved September 2016 from http://wtgrantfoundation.org/improving-use-research-evidence.

---

[1] See, e.g., National Research Council (2007); The America COMPETES Act of 2007, P.L. 1001-69; Members of the 2005 "Rising Above the Gathering Storm" Committee (2010); The America COMPETES Reauthorization Act of 2010, P.L. 111-358; Orszag and Holdren (2009); Donovan and Holdren (2015), National Science Board (2015).

[2] The "Scientifically Based Education Research, Statistics, Evaluation, and Information Act of 2000," introduced in the 106th Congress as H.R. 4875.

[3] This study was prepared under contract to the President's Commission on School Finance, charged by then President Richard M. Nixon with studying and reporting on "future revenue needs and resources of the nation's public and non-public elementary and secondary schools"; see Executive Order 11513 of March 3, 1970, included as an appendix to the Commission's 1971 progress report (U.S. President's Commission on School Finance, 1971).

[4] The full definition of scientifically based research in the Act indicates that such research: "employs systematic, empirical methods that draw on observation or experiment; involves rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions drawn; relies on measurements or observational methods that provide reliable and valid data across evaluators and observers, across multiple measurements and observations, and across studies by the same or different investigators; is evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random-assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition controls; ensures that experimental studies are presented in sufficient detail and clarity to allow for replication or, at a minimum, offer the opportunity to build systematically on their findings; and has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objectives, and scientific review."

[5] An example here is the *Common Guidelines for Education Research and Development* issued by NSF and IES in 2013.

[6]This assumes, of course, that scientific, technical quality is understood to be situation-specific, driven by and judged in light of the research questions of interest and contributions to knowledge investigators seek to make—points underscored, among other places, in *Scientific Research in Education* (NRC, 2002) and the *Common Guidelines for Education Research and Development* (NSF & IES, 2013).

[7] See http://cdn.carnegiefoundation.org/wp-content/uploads/2014/09/bryk-gomez_building-nics-education.pdf/ For a description of the Consortium's model for the role of research in supporting school reform, including the role the Consortium's Steering Committee plays in institutionalizing stakeholder consultation, see Roderick, Easton, and Bender Sebring (2009). For a discussion of the development and evolution of the networked improvement communities construct, see Bryk and Gomez (2008); Bryk (2009); and Bryk, Gomez, and Grunow (2010).

[8] See http://wtgrantfoundation.org/focusareas.

[9] See http://www.spencer.org/research-practice-partnership-grants.

[10] For a brief description of the Collaboratory (funded by NSF award # 1238253) see the project abstract at http://www.nsf.gov/awardsearch/showAward?AWD_ID=1238253&HistoricalAwards=false; for additional information see http://researchandpractice.org/vision/.

[11] For a brief description of the Center (funded by IES award # R305C140008) see the grant summary at http://ies.ed.gov/funding/grantsearch/details.asp?ID=1466; for additional information see http://www.ncrpp.org/.

[12] See, e.g., McDonald, Keesler, Kauffman, and Schneider (2006); Levin (2013).

[13] See, e.g., requirements for describing resources available and required to disseminate efficacy/replication, effectiveness, and measurement education research studies supported by the Institute of Education Sciences (U.S. Department of Education, 2015b), and guidelines for preparing plans for data management and sharing the products of research for proposals to the National Science Foundation (National Science Foundation, 2014).

[14] The What Works Clearinghouse is online at http://ies.ed.gov/ncee/wwc/. The "Find What Works!" interface can be used to access and compare evidence addressing school or district needs (see http://ies.ed.gov/ncee/wwc/findwhatworks.aspx).

[15] For more information on the Doing What Works library, see http://dwwlibrary.wested.org.

[16] See, e.g., Claridge and Fabian (2005); Soydan and Palinkas (2014); Beyea and Slattery (2013); U.S. Department of Justice, National Institute of Corrections (2013).

[17] Later that year he elaborated his argument in an editorial calling for better benchmarks to gauge how much a nation should spend on science (see Marburger, 2005b). These calls are now credited with establishing the science of science policy movement.

[18] NSF's Directorate for Social, Behavioral, and Economic Sciences (SBE) allocated $2.60 million in FY 2006 "to develop the data, tools, and knowledge needed to foster a new science of science policy" (NSF, 2006: 154) and the FY 2007 budget request to Congress identified "work to build a science of science policy and better science metrics through research in and across core programs" as a key priority for SBE's Division of Social and Economic Sciences. Within the year, NSF had developed plans through the new SciSIP program to support "fundamental research that leads to improved and expanded science metrics, datasets, and analytic tools from which researchers and policymakers may assess the impacts and improve their understanding of the dynamics" of the U.S. science and engineering enterprise (NSF, 2007: SBE-4—SBE-16). For more on the developing field of Science of Science Policy see e.g., Fealing, Lane, Marburger, and Shipp (2011); the Center for the Science of Science and Innovation Policy at the American Institutes for Research (http://cssip.org/); and the science of science policy website at http://www.scienceofsciencepolicy.net/.

[19] Notable here is the Star Metrics® (Science and Technology for America's Reinvestment—Measuring the EffecTs of Research on Innovation, Competitiveness, and Science) project, led by NSF, OSTP, and the National Institutes of Health. Building on scholarly research (including that supported by SciSIP), the Star Metrics initiative is ultimately expected to comprise "a broad collaboration of federal science and technology funding agencies with a shared vision of developing data infrastructures and products to support evidence-based analyses of the impact of science and technology investment" (Star Metrics®, 2015). Early, important successes of Star Metrics® and related initiatives (e.g., the Universities: Measuring the Impacts of Research on Innovation, Competitiveness, and Science [UMETRICS] initiative) are well-documented; see, e.g., National Research Council (2014a), Committee on Institutional Cooperation (2015).

[20] Examples that extend beyond the U.S. context include the U.K. Research Excellence Framework (REF), a "system for assessing the quality of research in UK higher education institutions" (http://www.ref.ac.uk/) and the U.K. Economic and Social Research Council's (ESRC's) impact reporting requirements via Researchfish. Both rely, in part, on tracking non-academic 'use' of research findings as a mechanism for gauging impact. For the 2014 REF, RAND Europe's *ImpactFinder* tool, initially developed to map impacts of studies supported by the Arthritis Research Campaign, was used by seven UK higher education institutions to identify likely impacts of research initiatives (see http://www.rand.org/randeurope/research/projects/impactfinder.html and http://www.rand.org/capabilities/solutions/assessing-and-articulating-the-wider-benefits-of-research.html). The ESRC requires grantees to file narrative impact reports which, among other things, provide "a summary of how the findings from your research grant are being used by the public, private or third/voluntary sectors, and elsewhere" (http://www.esrc.ac.uk/files/research/evaluation-and-impact/narrative-impact-guidance/). Researchfish is used by over 100 research organizations and funders in Europe and North America to report research outcomes and track impact (see http://www.researchfish.com/).

[21] See e.g., Bazerman and Moore (2012); COSEPUP (1986); Gigerenzer and Goldstein (1996); Gino (2013); Gino, Brooks, and Schweitzer (2012); Gladwell (2005); Heath and Heath (2013); Kahneman (2011, 1973); Kahneman and Tversky (2000); Kahneman, Slovic, and Tversky (1982); Manski (2005, 2007); Mauboussin (2009); Newell & Simon (1972); Russo and Schoemaker (1989); Simon (1972); Tallman, Leik, Gray, & Stafford (1993); Werthiemer (1922, 1959).

[22] See e.g., Cobb and Elder (1983); Cobb and Ross (1997); Cobb, Ross, and Ross (1976); Davis, Nutley, and Walter (2005); Downs (1972); Kingdon (1995); Meagher (2013); Nowotny, Scott, and Gibbons (2001); Peters (2007); Schattschneider (1960); and Tseng (2012).

[23] See, e.g., Backer (1991), Caplan (1979), Cousins and Leithwood (1993), Neilson (2001), and Nutley, Walter, Davies (2007) on research use in public services and public policymaking; Tetlock (2005) on political decision making; Simon (1960) and Bazerman and Moore (2012) on managerial decision making. On research and knowledge utilization more generally, see Davis and Nutley (2008), Knott and Wildavsky (1980), Larsen (1980), Estabrooks et al (2008), Rich (1997), and Weiss (1977, 1979, 1988, 1995).

[24] See e.g., Barstow, Dunleavy, and Tinkler (2014), Davis, Nutley, and Walter (2005), LSE Public Policy Group (2008), Lynd (1939), and Mollett, Moran, and Dunleavy (2011).

[25] See e.g., ESRC (1994), Pettigrew (2011), and Starkey and Madan (2001).

[26] See e.g., Coburn, Honig, and Stein (2009), Corcoran (2003), Cousins and Leithwood (1993), Eidell and Kitchel (1968), Finnigan and Daly (2014), Hood (2002), Lehming and Kane (1981), NEKIA (2005), and Paisley and Butler (1983).

[27] The William T. Grant Foundation launched its research evidence initiative in 2008; at that time the Foundation was particularly interested in "increase[ing] understanding of how research is acquired, understood, and used, as well as the circumstances that shape its use in decision making" (William T. Grant Foundation, 2015: 2). Recently, the Foundation announced a somewhat new direction for its research evidence funding; beginning with the 2016 competition the Foundation is shifting its focus "from understanding how and under what conditions research is used to understanding *how to create those conditions*" (*Ibid.*, p. 3; emphasis in the original).

[28] The two IES-funded Centers, funded as the result of IES FY14 and FY15 competitions, are the National Center for Research in Policy and Practice (NCRPP) at the University of Colorado (IES award #R305C140008, see http://ncrpp.org/) and the Center for Research Use in Education (CRUE) at the University of Delaware (IES award # R305C150017, see http://ies.ed.gov/funding/grantsearch/details.asp?ID=1641).

[29] These include portions of the U.S. Department of Education Institute of Education Sciences' (IES) RFA pertinent to mathematics and science learning (a single request for applications covers all education research grants supported by IES), and program descriptions and solicitations or announcements for five National Science Foundation Directorate for Education and Human Resources programs that support STEM education research in K-12 settings and/or for students of corresponding ages.

[30] An expanded treatment of research's enlightenment function is provided in an earlier *Policy Analysis* article (Weiss, 1977).

[31] See, e.g., McCombs and Shaw (1972); McCombs (2004).

[32] Based on the application of a "parameter-free measure that quantifies the extent to which a specific paper can be considered an SB [Sleeping Beauty] . . . to 22 million scientific papers published in all disciplines of natural and social sciences over a time span longer than a century", Ke and colleagues conclude ". . . papers whose citation histories are characterized by long dormant periods followed by fast growths are not exceptional outliers, but simply the extreme cases in very heterogeneous but otherwise continuous distributions" (Ke, Ferrara, Radicchi, & Flammini, 2015: 7426 & 7431).

[33] This volume presents the response of Director of the Office of Scientific Research and Development Vannevar Bush to President Roosevelt's request that he explore the peacetime implications of a war time (WWII) experience "coordinating scientific research and . . . applying existing scientific knowledge to the solution of…technical problems".

[34] For example, when the National Science Foundation filed its first annual performance report under the Government Performance and Results Act of 1993 (GPRA), then Director Rita Colwell observed that outcome goals for NSF investments "focus on the long-term results of NSF's grants for research and education in science and engineering" (NSF, 2000: Director's Statement). The report clarified that this included "[c]onnections between discoveries and their use in service to society" (*Ibid.*, p. 4).

[35] One approach involves post hoc construction (often by external evaluators or funders) of illustrative vignettes. Other approaches look to investigators to describe the uses that have been made of their research findings; an example here is the UK Economic and Social Research Council's (ESRC's) requirement that awardees report not only outcomes, but (one year post conclusion of their grants) narrative impact statements including a summary of how research grant findings are being used by the public, private, or third/voluntary sectors" (see ESRC's instructions to awardees for completing the Narrative Impact Report, accessible via http://www.esrc.ac.uk/funding/guidance-for-grant-holders/reporting-guidance/).

[36] For more on the construction and functions of logic models in program management and evaluation see e.g., Frechtling (2007), McLaughlin and Jordan (1999), Miller, Simeone, and Carnevale (2001), Taylor-Powell, Jones, and Henert (2003), and W.K. Kellogg Foundation (2004).

[37] Similarly, key assumptions are typically summarized separately from the chain of action drawn in the model.

[38] For more on the limitations and effective uses of logic models, see e.g., Funnell and Rogers (2011), Hummelbrunner (2010), and Renger, Wood, Williamson and Krapp (2011).

[39] Evaluability assessments frequently establish readiness-for-evaluation and clarify program theory, the purposes of conducting an evaluation/intended uses of insights from evaluation, and methodologies appropriate to address evaluative questions; see e.g., Smith (1989), Trevisan (2007); Trevisan and Huang (2003), Wholey (2004).

[40] A Center for Research on the Utilization of Scientific Knowledge was established at the University of Michigan's Institute for Social Research in 1964 to "address the failure of organizations and individuals to make use of scientific knowledge" (Institute for Social Research, 2015). Specific Center goals included examining "how individuals and groups spread and use knowledge", analyzing "approaches to encourage knowledge utilization", and considering "the broader moral issues involved". CRUSK's advisory committee was headed by Ronald Lippitt; its first director was Floyd Mann. Twenty some years after it was established, CRUSK was disbanded. Source: Institute for Social Research (2015).