Chapter 48
Trends in US Government-Funded Multisite K—12 Science Program Evaluation*

Frances Lawrenz and Christopher David Desjardins

The importance of science education in the USA's economic and security interests has been highlighted in a number of national reports. A recent report from the US National Academies cited the significance of science education in maintaining the USA's competitive edge in the world economy (2007). The National Science Board (2007) addressed a declivity in the career choice of engineering as well as a general weakness in the K—12 science, technology, engineering, and mathematics (STEM) curriculum, citing that engineering is the key to an innovative, technological society. A pervasive example of this concern is the Elementary and Secondary Education Act of 1965 and its reauthorization as the No Child Left Behind (NCLB) Act of 2001.

Evaluation of government programs to enhance science education is an ongoing process greatly affected by the political environment as well as the government agency providing the program. Examination of US federal science programs and their evaluation over time highlights the effects of the changing context and its attendant values on science education.

In this chapter, we present the history of federally funded science education programs and their evaluation by examining selected US government agencies involved in science education. We begin with a description of the science-education-oriented federal agencies followed by a definition of multisite science program evaluation. We continue with a history of the multisite science education programs and evaluations in the National Science Foundation (NSF), National Aeronautics and Space Administration (NASA), and National Oceanic and Atmospheric Administration (NOAA). We then relate these histories to the changing political contexts and changes in evaluation research and theory.

*This work was partially supported by IES Award # R305C050059 Interdisciplinary Education Sciences Training Program University of Minnesota PRF# 473473.

F. Lawrenz (✉) • C.D. Desjardins
Educational Psychology, University of Minnesota, Minneapolis, MN, USA
e-mail: lawrenz@umn.edu; desja004@umn.edu

B.J. Fraser et al. (eds.), *Second International Handbook of Science Education*,
Springer International Handbooks of Education 24, DOI 10.1007/978-1-4020-9041-7_48,

US Federal Agencies Providing K—12 Science Education

In 2005, 90% of the $536 billion spent on education came from state and local funding, with only 10% provided by the federal government (U.S. Department of Education [DoEd] 2006). Although the majority of funding for K—12 education comes from state and local sources, the federal government plays an important role in science education in two ways: (1) the federal government passes legislation that affects federal funding, for example, NLCB, and (2) by providing funds for federal agencies to use for education. The role of federal agencies in science education has been reviewed by two federal cross-agency panels since 1993: the Federal Coordinating Council on Science, Engineering, and Technology (FCCSET) (Federal Coordinating Council for Science, Engineering and Technology 1993) and the Academic Competitiveness Count (ACC) (DoEd 2007). Both found that federal agencies have an important role in K—12 science education and stressed the need for collaboration and coordination. In testimony of the widespread importance of science education to the federal government, the Academic Competitiveness Council Report (DoEd 2007) discloses that there are currently 12 federal agencies that provide funding for STEM education. Eight of these agencies provide funds specifically for K—12 STEM programs. The report goes on to say that in 2006 federal agencies spent $3.1 billion on STEM education, $574 million (18%) of which supported K—12 science education programs.

Agencies and departments such as NASA, Department of Energy, and NOAA, are designed to provide science services to the nation through such things as the space program, the national energy laboratories, weather mapping, etc. These mission agencies have direct access to scientists and cutting-edge science, but not necessarily educational expertise. They are usually interested in science education in an effort to keep students in the STEM pipeline to provide a strong workforce and support general scientific literacy. They engage in substantial outreach activities, mostly in the form of science education programs (e.g., NOAA's B-WET program). Other mission agencies provide mostly direct services and may have some outreach activities related to science education (e.g., the National Park Service's visitor centers and programs).

The two most important agencies in science education, providing about 85% of the federal funds in 2006, are the (DoEd) and NSF. DoEd's K—12 STEM-specific education budget represents less than 1% of its total 2006 investment (National Research Council of the National Academies [NRC] 2008). Most of the funding presently goes to the Mathematics and Science Partnership (MSP) program, a formula grant program whose mission is to develop rigorous STEM curricula in K—12, distance learning programs, and incentives to entice STEM majors into the teaching profession. MSP-type programs were formerly funded at a higher level under the older Eisenhower program. DoEd supports research through the Institute of Education Sciences (IES) that was established in 2002.

The Education and Human Resources Directorate (EHR) at NSF provides funding for science education through its limited-term grants for educational research, innovative curriculum development and pedagogy, teacher professional development, education programs and activities, and other education initiatives. EHR's budget was about $797 million in 2006, of which around $22 million (30%) supported K—12 science education. Most recently, EHR has increased attention to research on learning and teaching, and has reorganized its research grant programs related to teaching and learning into a single division. Other directorates at NSF also support education initiatives, such as the Directorate for Engineering.

Evaluation of Programs

The Joint Committee on Standards for Educational Evaluation (1994) has defined evaluation as the systematic investigation of the worth or merit of an object. Scriven (1991) suggests that evaluation also includes the identification of relevant standards of worth. These terms (merit, worth, standards) highlight the intimate connection of evaluation with the value systems of the people commissioning, conducting, participating in, and receiving the evaluation. Because it differs in intent, evaluation can be considered distinct from research (Weiss 1988).

All federal agencies are subject to evaluation by the Office of Management and Budget (OMB). The reporting to OMB has taken a variety of forms over the years, most recently as the Government Performance and Results Act of 1993 (GPRA) and the Performance Assessment Rating Tool (PART, available at http://www.whitehouse.gov/omb/part/fy2007/2007_guidance_final.pdf). This tool requires each agency to demonstrate how performance of their programs will be measured. Performance measures can be both long term and annual, and must reflect program goals and include verifiable data collected through reliable research methods. Coupled with PART is the work of the ACC which examined the overlap among federal groups working on science education. As a result, ACC recommended types of designs to use in conducting evaluations of or research about science education programs (DoEd 2007). The granting agencies use a Committee of Visitors process where a team of field-based experts comes into the agency, reviews the quality of the funded proposals, and produces a report. Many reviews have been conducted by the National Academies which was given the authority to advise the USA on scientific and technical matters in 1863. The National Research Council (NRC) was organized by the National Academies in 1916 to associate the broad community of science and technology with the Academy's purposes and has become the principal operating agency. A final method of evaluation is for the agency to contract with an external evaluator to assess a particular program. See, for example, the externally contracted final evaluation report of the Local Systemic Initiative by Banilower et al. (2006).

The history of federally funded science programs and evaluations is one of differing but repeated emphases. These emphases mirror societies' expectations of science programs in terms of curriculum, teacher professional development, student assessment, perceived locus of change, and national leadership and requirements.

Perhaps one of the first implementations of evaluation in the USA was Joseph Rice's comparative study of spelling performance (1898). The next landmark was the Eight Year Study by Tyler and Smith (1942). A 1994 review of science education assessment (Doran et al. 1994) revealed that the 1960s laid the groundwork for present-day science education program evaluation. US federal program evaluation became widespread with the development of the National Assessment of Educational Progress (NAEP), the proliferation of Great Society social programs in the mid-1960s, and the passing of the Elementary and Secondary Education Act of 1965 that mandated evaluations of Title I and Title III education programs (Fitzpatrick et al. 2003).

Questioning the non-utilization and underutilization of evaluations began during the 1970s as evaluators became increasingly concerned about the utility of their evaluations. Such concerns arose in light of economic uncertainty due to recessions and inflation, perceived failures of many Great Society programs in conquering societal ills, and the Watergate scandal that led to great mistrust of the federal government. As quoted by then chairman of the Committee on Labor and Human Resources in the foreword to a volume entitled *Evaluation in Legislation*, “politics has gone from the age of ‘Camelot’ when all things were possible to the age of ‘Watergate’ where all things are suspect” (Williams 1979, p. 8).

During the 1980s, maximizing the impact of evaluation became increasingly important. Arguably, three factors contributed to this new emphasis. First, the 1980 election of Ronald Reagan, bringing a fiscally conservative political stance, presented both challenges and opportunities for evaluators. Second was the movement toward professionalization of the field of evaluation. Early steps of this movement included the appointment by a dozen leading educational organizations of a committee of educational evaluators and researchers in 1975 and the subsequent publication of the Joint Committee’s *Standards for Evaluations of Educational Programs, Projects, and Materials* in 1981. Third was the advancement of social science methodology. Social science researchers began to value integrative reviews and meta-analyses as forms of research that were complementary and not just secondary to individual research studies. Over time, collaborative and participatory evaluation models began to arise. These models involved planning for use early in an evaluation and involving intended users in the process to increase the effectiveness and tangibility of the process and its findings.

Recently, there have been several trends in evaluation. One is the revision of the Program Evaluation Standards (Joint Committee WMU) and the development of the Guiding Principles for Evaluation (AEA web site), making them more compatible with changing evaluation needs. Another is the emphasis on including diverse perspectives in evaluation planning (Greene et al. 2006) or the culturally responsive approach championed by Mertens (2005). Additionally, as mentioned above, the US ACC (DoEd 2007) advocates a heavily quantitative approach. Lastly, a strong emphasis has been placed on evaluation capacity building and participant involvement in the evaluation process, especially as it relates to increasing evaluation use and influence.

History of NSF K—12 Science Education Programs and Evaluations

As the main federal science education program funder, NSF is a primary example of the effect of history on science programs and their evaluation. NSF's approach to science education programs has been somewhat cyclical. After Sputnik, NSF focused on improving science education through teacher professional development and the construction of new curricula to help win the "race for space." Additionally, the National Defense Student Loan was created to help encourage people to become science teachers by forgiving a portion of the loan for each year spent as a teacher in the program. Evaluation concentrated on the scientific accuracy and effectiveness of these curricula and the newly prepared teachers in helping students learn science.

During the Vietnam era, significant distrust of the government caused NSF programming to switch from large-scale to local programs. These programs were often summer institutes, designed to enhance teacher understanding of science and mathematics and teacher pedagogical skills. Evaluations focused on perceived quality and were individualized to the needs of the programs and their stakeholders (Lawrenz 2007). After continuing for some time, the late 1980s saw an increase in large-scale programs with the Systemic Initiatives. The Systemics included statewide, urban, rural, and local school district programs. Evaluation was much more complex and assessed how to change cultures as well as interactions and the results those changes might produce. This produced the beginnings of national databases to track status information and centralized or pooled approaches to conducting evaluations. In addition, it led to the realization that this sort of evaluation takes a good deal of time and money. Large-scale programs showed up again in the late 1990s with MSP and the Centers for Learning and Teaching. Evaluation was complex with a heavy emphasis on accountability and direct ties to state-based testing systems (Lawrenz 2007). Measures of organizational change and promotion of interaction were developed. Furthermore, several research, evaluation, and technical assistance projects were funded to assist the partnerships with their evaluations (Lawrenz 2007).

Most recently, NSF is emphasizing the research aspects of its programming and is interested in funding transformative ideas. In-service teacher master degree programs have returned as the teacher institute component of the MSP program. The preservice teacher scholarship program idea has resurfaced in the form of the Noyce program. Science program evaluation has moved toward more randomized designs and sophisticated regression-based modeling. Often, yearly achievement data required by the NCLB initiative, national study data such as Trends in International Mathematics, and Science Study (TIMSS) or national longitudinal studies are used.

History of NASA K—12 Science Education Programs and Evaluations

NASA has been in operation since 1958, directly after the launch of Sputnik in October, 1957. NASA's role in K—12 science education is closely linked to and guided by its core scientific, engineering, and exploration missions. NASA provides about 4% of the federally sponsored K—12 education.

NASA has been involved in education since its early years with the Aerospace Education Services Project (AESP) established in 1962. The bulk of the K—12 science education activities are in the Office of Education and the Science Mission Directorate (SMD). Each accounts for about 50% of the agency's total K—12 funding. The SMD devotes a percentage of funds, connected with each major science mission to education activities. The amount of funding for education has been decreasing; for example, the budget for the Office of Education decreased from $230 million in 2003 to $153 million in 2007.

The mechanisms by which these two entities functioned have changed over the years. Prior to 1992, programming was quite independent and K—12 education projects tended to evolve as a diverse portfolio of often disconnected activities. In 1992, however, NASA established its first agency-wide education strategy. The objective for K—12 then, which remains much the same today, was to use NASA's mission to enhance the content, knowledge, skill, and experience of teachers; to capture the interest of students; and to channel that interest into related career paths through the demonstration of the application of science, mathematics, technology, and related subject matter. In 1996, the implementation plan emphasized scientists working in high-leverage partnerships with educators. Most of the education projects in the science and technology enterprises were located in the Office of Space Science (OSS) and the Office of Earth Science (OES). OSS programs generally involved grants for scientists working with educators to provide educational experiences. The OES projects were more traditional in terms of providing curriculum and professional development. The NASA centers played a central role through their education coordinators and the development of center-specific projects. Education coordinators promoted extensive outreach and engagement with local schools and informal science education services.

Recently, NASA programming has been experiencing administrative change due to political pressure. For example, since 2000, NASA educational programs have been organized to align to three different agency-wide strategic plans. In 2003, there was an internal review of the 48 K—12 programs and only those perceived as effective were continued. The OSS and OES were merged into a new directorate that includes the majority of the mission-oriented educational programs. Most recently, all K—12 projects are to focus primarily on attracting and retaining students in science disciplines through engagement and educational opportunities. K—12 projects are divided into four major categories, educator professional development of less than 2 days, educator professional development of more than 2 days, curricular support resources, and student involvement. Coordination and management of the various programs has been distributed to the various centers.

Only a limited number of evaluations have been conducted on these programs. Only three of the programs, the NASA Explorer School, the Aerospace Education Services Project, and the Science Engineering Mathematics and Aerospace Academy, have been substantially evaluated. As part of the NRC (2008) report, a detailed critique of the available evaluations of the NASA programs was also prepared. The critique provided information about the methods and the results of NASA evaluations.

All evaluations reported on how the program was operating and how that operation fit within NASA goals. All provided recommendations as to how the program might be improved or changed. Most provided a good deal of information about how the participants in the program felt about the program. Overall, they provided very interesting descriptive information about the programs from the perspectives of those involved. However, the samples used to gather evaluation information were often convenience samples; meaning the people used were those from whom data were easy to obtain. Results yielded perceptions that were overwhelmingly positive. There were only a very few small attempts at comparative studies and these were flawed by selection bias; one group was likely to have been different from the other at the start.

History of NOAA Science Program Evaluations

Although NOAA was first formed in 1970, the agencies that came together at that time are among the oldest in the federal government. The agencies included the US Coast and Geodetic Survey formed in 1807, the Weather Bureau formed in 1870, and the Bureau of Commercial Fisheries formed in 1871. As the USA's leading oceanic and atmospheric science and service agency, NOAA has the responsibility to increase its coordination and collaboration within the ocean, coastal, Great Lakes, weather, climate science, and education communities. The administration has had a federally mandated educational mission since at least 1966 with the passing of the National Sea Grant College and Program Act. Most recently in 2007, NOAA's role in earth system science education was solidified by the America Competes Act. This legislation provided NOAA a mandate to advance its educational efforts, and engage a broader community of partners in creating an environmentally literate society as well as a viable workforce of scientists, managers, and administrators in support of a sustainable future (National Oceanic and Atmospheric Administration [NOAA] 2008b). The high interest at NOAA for evaluation is exemplified by the first outcome listed on its Education Strategic Plan ("evaluation and research for effective programs" (NOAA 2008b).

NOAA's organizational chart shows its Office of Education as reporting separately from the six operating branches. Both the operating branches and the Office of Education provide science education programs. The Office of Education and the agency-wide Education Council were formed in 2003 as part of the agency's commitment to environmental literacy as a cross-cutting priority. Programs are provided in both formal (K—12 schools, colleges, etc.) and informal settings (after school programs, museums, etc.) for teachers, students, and the general public of all ages. NOAA partners with other agencies and professional groups to help develop its educational programs. For example, the Essential Principles of Ocean Literacy (National Geographic Society 2006) and Essential Principles of Climate Literacy (NOAA 2008a) were developed to help guide educational efforts. The Office of Education operates an Environmental Literacy grants program which began in 2005. As of 2007, this program provided $1.6 million for Science on a Sphere projects in science museums and centers as well as $6.8 million to 15 free choice and K—12 formal education programs. The 2006 budget showed the following breakdown of education and outreach areas: Climate (2%); Weather and Water (2%); Ecosystems (43%); Commerce and Transportation (5%); and Mission Support (48%).

NOAA has a broad array of science education programs and these programs have been affected by the political environment. NOAA has responded to the differing national science education agendas by providing ocean education, environmental education, and most recently, climate change education. Much work has been done to counteract the perceived lack of emphasis on earth sciences in the National Science Education Standards. Some programs directly focused on K—12 science education are Sea Grant, Ocean Exploration, Teacher at Sea, Storm Ready/Tsunami Ready, Bay Watershed Education and Training Programs, and Jason.

As one of NOAA's longest funded educationally related programs, the Sea Grant program has been the most evaluated. In fact, in addition to a comprehensive regular evaluation procedure involving external review and rankings, the program was twice evaluated by the National Academies. The first report in 1994, A Review of the NOAA National Sea Grant College Program, suggested changes to the comprehensive regular evaluation review procedures (NOAA 1994). In 2006, a second evaluation (NOAA 2006) examined the effects of the 1994 report in Evaluation of the Sea Grant Program Review Process. Almost all of the NOAA educational programs are evaluated in some way. Overall the evaluations are much like those described for NASA, although the NOAA evaluations tend to be more quantitatively oriented.

In 2007, the National Academies were requested by the NOAA Office of Education to review the NOAA education programs. This 3-year review will result in a comprehensive report addressing the role of NOAA, the appropriateness of its goals and objectives, the effectiveness of the educational programs, the composition of its education portfolio, and the quality of the evaluations of its programs. Including evaluation as one of the major questions for the review highlights the importance of evaluation and accountability within the agency.

Implications

The US federal government plays an important role in science education, even though its total contribution to the K—12 education budget is relatively small. The agencies, especially NSF, are viewed as providing a leadership role in what is important for science education. Mission agencies such as NASA and NOAA also play an important role in promoting their specific areas of science education. All agencies provide their programs as incentives for schools to participate; however, schools are not required to participate. Even federally mandated programs such as NCLB are voluntary with the withdrawal of federal support used as an impetus to participate.

Other governments around the world have similarities and differences in terms of the way they participate with K—12 level science education. For example, in Singapore, science education in the grades corresponding to the US K—12 system is nationally supported through the Ministry of Education. The Ministry also supports the National Institute of Education as its research arm, much like the US Institute for Educational Sciences. Although Singapore's government does support science research agencies like the USA's National Institutes of Health (NIH), those agencies are not engaged in K—12 science education. Singapore has a national curriculum, one part of which is science.

As another example, in Australia a national curriculum is just being developed (beginning in 2009) whereas in the past each of the states had developed its own curriculum, much like in the USA. Until very recently, most of the funding for the equivalent to K—12 education flowed through the Australian federal government into the states. The states functioned mostly independently, although the federal government made suggestions as to how the money should be used. This is also similar to the US DoEd's flow through block grants to the states, although the withholding of federal money is enough of a stick that most states in the US conform to federal recommendations.

This review has documented the types of science programming and concomitantly, program evaluation experienced in three federal agencies. These show that federally sponsored science programs and their evaluations are closely tied to political agendas and contexts. NSF science education programming emphases have been somewhat cyclical, oscillating from large to local programs and from implementation to research as public opinion of the government and government priorities have changed. NASA science education programming emphases have been responsive to public opinion about space programs and science and engineering as appropriate career paths. NOAA science education programming has reflected the public interest in the environment, especially oceans and weather. As the emphases in science education programs differ, the evaluations differ in terms of what they value and how they measure valued outcomes. In recent years, there has been more emphasis on gathering summative data for accountability and consequently there has been much less emphasis on formative evaluations across all federal agencies. Despite recent calls for more comparative studies to assess accountability, programs or even projects within agencies are seldom compared, much less programs compared across agencies. Despite this proclivity, there have been attempts to look across agencies (e.g., ACC and FCCSET). The US government agencies tend to pass along their own requirements for evaluation (e.g., GPRA and PART) to the programs with which they work. National interest in the goals of the agencies appears to govern the type of programming more than the results of evaluations. For example, climate change is an important recent topic and programs on climate change will be supported, regardless of evaluation data.

A consistent and increasingly more salient goal across the agencies has been expanding the diversity of people engaged in science and science education. For example, many directorates at NSF fund programs to attract underrepresented groups, and NOAA has a diversity council to address these issues. Similarly, NASA has several related programs including the Introduce a Girl to Engineering Day. Not only are science programs provided to explicitly address issues of underrepresentation, but also to attend to cultural responsiveness (e.g., Mertens and Hopson 2006). The recent revision of the NSF's *User-Friendly Handbook for Project Evaluation* includes a chapter on culturally responsive evaluation (Frechtling 2002).

It is clear through the many reports, acts, and laws surrounding science education that the US federal government is very interested in science education. Its rationale for that importance changes from strategic military needs, to prestige, to economic advantage. However, the call for improvement is consistent. The involvement of the different agencies makes the response somewhat ad hoc, but concurrently responsive to individual needs and interests. It is unlikely that most K—12 science educators are aware of the plethora of science education experiences that are available. Much of the programming is accessible in limited geographical areas or to select people through word of mouth. This is truly unfortunate. Science educators should call for more coordination of the federal programming and more efficient information dissemination techniques. A coordinated program with each agency contributing what it does best would likely be more efficient than the existing independent programming.

If a federally coordinated program existed for science education, evaluation could be conducted on a larger scale and produce more generalizable results. In turn, this would help to increase the effectiveness of the programming. If such evaluations were possible, science educators should advocate for diversity of perspectives and methods, as well as high quality and rigor. Critical and interpretive methods (e.g., Coghlan et al. 2003) should be balanced with more positivistic approaches (DoEd 2007). It would also be important to evaluate the effectiveness of the different evaluation methods being used to examine science education. As a result, the methods themselves could be improved (Burkhardt and Schoenfeld 2003). Finally, although there has been work identifying the essential competencies required of an evaluator, there is no clear indication of what skills might be explicitly needed for science program evaluation (Stevahn et al. 2005).

References

Banilower, E. R., Rosenberg, S.L, & Weiss, I. R. (2006). Local systemic change through teacher enhancement, 2003-06 Cross-Site Report. Retrieved October 21, 2008 from http://www.horizon-research.com/LSC/news/cross_site/06cross_site/cross-site06.pdf

Burkhardt, H., & Schoenfeld, A. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, *32*(9), 3–14.

Coghlan, A. T., Preskill, H., & Catsambas, T. T. (2003). Using appreciative inquiry in evaluation. In H. Preskill and A. T. Coghlan (Eds.), *New directions in evaluation* (pp. 5–22). San Francisco: Jossey-Bass.

Doran, R., Lawrenz, F., & Helgeson, S. (1994). Research assessment in science. In D. Gabel (Ed.), *Handbook for research teaching and learning* (pp. 388–442). New York: Macmillan Publishing Company.

Elementary and Secondary Act of 1965. Public Law 89-10, 89th Congress, 1st Session, April 11, 1965.

Federal Coordinating Council for Science, Engineering and Technology. (1993). *The federal investment in science, mathematics, engineering, and technology education: Where now? What next? Sourcebook.* Report of the expert panel for the review of federal education programs in science, mathematics, engineering, and technology (ERIC Document Reproduction Service No. ED366502.) Retrieved October 21, 2008 from http://www.eric.edu.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/41/fc.pdf

Fitzpatrick, J. L., Worthen, B. R., & Sanders, J. R. (2003). *Program evaluation: Alternative approaches and practical guidelines* (3rd ed.). White Plains, NY: Longman.

Frechtling, J. (2002). *The 2002 user-friendly handbook for project evaluations.* Prepared under contract: REC99-12175. National Science Foundation: Directorate for Education and Human Resources, Division of Research, Evaluation and Communication.

Government Performance Results Act of 1993. Public Law, 103-62, 103rd Congress, 1st Session, January 5, 1993.

Greene, J. C., DeStefano, L., Burgon, H., & Hall, J. (2006). Advancing evaluation of STEM efforts through attention to diversity and culture. In D. Huffman & F. Lawrenz (Eds.), *Critical issues in STEM evaluation* (pp. 53–72). San Francisco: Jossey-Bass.

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Lawrenz, F. (2007). Science program evaluation. In S. K. Abell and N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 943–964). Mahwah, NJ: Lawrence Erlbaum.

Mertens, D. M. (2005). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative and mixed methods* (2nd ed.). Thousand Oaks, CA: Sage.

Mertens, D. M., & Hopson, R. K. (2006). Advancing evaluation of STEM efforts through attention to diversity and culture. In D. Huffman & F. Lawrenz (Eds.), *Critical issues in STEM evaluation* (New Directions for Evaluation, No. 109) (pp. 35–51). San Francisco: Jossey-Bass.

National Geographic Society. (2006). *Essential principles of ocean literacy*. Retrieved October 21, 2008 from http://www.ngsednet.org/community/resource_uploads/OceanLitBrochure.pdf

National Oceanic and Atmospheric Administration. (1994). *A review of the NOAA national sea grant college program*. Washington, DC: National Academy Press.

National Oceanic and Atmospheric Administration. (2006). *Evaluation of the sea grant program review process*. Washington, DC: National Academy Press.

National Oceanic and Atmospheric Administration. (2008a). *Essential principles of climate literacy*. Retrieved October 21, 2008 from http://www.climate.noaa.gov/education/pdfs/climate_literacy_poster-final.pdf

National Oceanic and Atmospheric Administration. (2008b). *NOAA's education strategic plan (Draft)*. Retrieved October 20, 2008 from http://www.oesd.noaa.gov/08%20NOAA%20Education%20Strategic%20Plan%20Full%20Draft%207_14.pdf

National Science Board. (2007). *Moving forward to improve engineering education* (NSB-07-122). Washington, DC: Author.

National Research Council of the National Academies. (2008). *NASA's elementary and secondary education program*. Washington, DC: National Academy Press.

No Child Left Behind Act of 2001. Public Law 107-110. H.R. 1. 107th Congress, 2nd Session, January 8, 2002.

Rice, J. M. (1898). *The rational spelling book*. New York: American Book Company.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage Publications.

Stevahn, L., King, J., Ghere, G., & Minnema, J. (2005). Establishing essential competencies for program evaluators. *American Journal of Evaluation*, *26*(1), 43–59.

Tyler, R. W., & Smith, E. R. (1942). *Appraising and recording student progress*. New York: McGraw-Hill.

U.S. Department of Education. (2006). *The condition of education 2006* (U.S. Department of Education NCES 2006-071). Washington, DC: Ed Pubs.

U.S. Department of Education. (2007). *Report of the Academic Competitiveness Council*. Washington, DC: Author.

U.S. National Academies. (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: The National Academies Press.

Weiss, J. (1988). Quality education indicators in Canada: What will "stand on guard for thee"? *Studies in Educational Evaluation*, *14*, 65–74.

Williams, H. A., Jr. (1979). Foreword. In F. M. Zweig (Ed.), *Evaluation in legislation* (pp. 7–9). Beverly Hills, CA: Sage Publications.