

Steady Hand at the Wheel: Using an Incremental, Systematic Approach to Promote an Evaluation Culture in a Federal Organization

Yaw O. Agyeman, Ph.D., Lawrence Berkeley National Laboratory,

ABSTRACT

The U.S. Department of Energy (DOE) Office of Energy Efficiency and Renewable Energy (EERE) has gradually emerged as the leading face of evaluation across the entire Department of Energy, prompting the White House's Office of Management and Budget (OMB) to refer to EERE as exemplar in evaluation practice, despite a list of formidable odds. The relative success that EERE has had in moving the organization (comprised of 10 energy efficiency and renewable energy offices) towards a culture of evaluation belie a staunch and obdurate resistance born of the usual objections: 1) the difficulties of evaluating R&D and technology deployment programs; 2) the political repercussions if findings are negative; 3) the absence of any perceived benefits even if findings are positive; and 4) the absence of dedicated resources to advance an evaluation culture. We developed a strategic, incremental approach to overcome these and several, additional entrenched objections, sometimes through careful planning, sometimes through the stroke of fortune. Central to our approach has been: 1) careful targeting of sympathetic program leads, with whom we partnered to conduct some successful evaluations; 2) developing a message to leadership to sell evaluation as the useful tool for management that it is, and emphasizing the benefits accrued to offices that joined us to evaluate some of their investments; 3) developing accessible guides and standards to support the conduct of evaluation; 4) providing training opportunities to staff to develop their skills in evaluation; and 5) convincing upper management to make the use of evaluative information a performance assessment requirement. We discuss the progress made to date with our evaluation strategy, and our thinking on next steps in strengthening a lively, useful culture of evaluation in the organization.

Introduction

The purpose of this brief is to outline the path being charted by the evaluation lead in the United States Department of Energy's (DOE) Office of Energy Efficiency and Renewable Energy (EERE) to instate, promote and build a culture of evaluation within the organization. The effort to build a culture of evaluation has been painstaking, incremental, and long, endangered and buoyed, by turns, by crosswinds of favorable and unfavorable political winds. So many years on, the transformative effort remains, in many ways, still in its infancy, despite the undeniable progress made. We address some of the key issues, achievements, lessons learned, and the way forward.

Characteristics of an evaluative culture

An evaluative culture is defined by a set of general characteristics that, taken together, serve to ensure that the organization is efficient and effective. These characteristics, as described by Mayne (2008), include a constant engagement in:

- Self-reflection and self-examination, involving:

- Purposeful seeking of evidence on what is being achieved, such as through monitoring and evaluation
- Using results information to challenge and support program activities
- Valuing honesty, challenge and genuine dialogue
- Evidence-based learning, involving:
 - Making time to learn in a structured way
 - Learning from mistakes
 - Encouraging knowledge sharing
- Experimentation and change, involving:
 - Supporting risk taking, and
 - Seeking new ways of doing business when the evidence points to the need for such

In sum, the evaluative culture is one that purposefully seeks out information on how it is performing, both in terms of outputs and outcomes, in order to use that information to learn how to better manage and deliver its programs and services, and thereby improve its performance. Campbell (1969) framed the underlying concepts as a guide for a dynamic, experimenting society – concepts adapted subsequently by Trochim and Donnelly (2006) as a set of values for an evaluation culture. The promulgated values view an evaluation culture as one that is:

- *Action oriented* – actively seeking solutions to problems, and being willing to experiment
- *Teaching oriented* – promoting the relatedness of formal evaluation and everyday thought, accessible to all, encouraging experts to share their knowledge and embodying the concept of a “learning organization”
- *Diverse and inclusive, participatory, responsive and fundamentally non-hierarchical* – promoting the understanding that problems and the solutions devised for them are usually systemic, interconnected and linked to social issues
- *Humble and self-critical* – acknowledging the limitations of individual studies, recognizing that decisions require multiple inputs, and recognizing who has responsibility for decisions in democratic societies
- *Interdisciplinary* – purposefully moving beyond the tunnel focus of disciplines to learn from the lessons of various disciplines, and abdicating the confusion of jargon to reach a broader audience
- *Honest and truth seeking* – emphasizing accountability and scientific credibility, and upholding the goal of getting at the truth while honestly acknowledging that all scientific knowledge is revisable
- *Prospective and forward looking* – strategically anticipating where information will be needed and initiating low-cost monitoring information systems during program initiation
- *Fair, open, ethical, and democratic* – shifting away from private ownership and exclusive access to data, promoting open commentary and debate on the results of evaluations and providing formal opportunities for competitive review and response

Challenges to overcome in EERE

Obstacles to the establishment of an evaluative culture are inherent, to varying degrees, in all organizations, and tend to come in similar flavors, as we found for EERE. The compendium of challenges to the mere commissioning of an evaluation study, never mind establishing an evaluative culture, is long, and entrenched, and includes some familiar laments:

- It is too difficult to evaluate investments in research and development (R&D), and the deployment of technologies and processes emanating from these;
- The political pitfalls are too great, and potential repercussions of negative findings too high;
- There is little to no benefit to evaluations, even when results are positive;
- Program funds are already limited; it is better to focus allocation on program activities than on evaluation; and
- We already know that our programs are effective; we do not need to invest in evaluations to tell us what we already know.

Evaluation professionals might find these refrains to be familiar, and might offer point-by-point counter-arguments to favor the merits of evaluative practice. But the potency of the resistance of program staff has often been stronger, because their complaints have frequently proven true, and well founded, drawn from a long-standing, entrenched organizational history filled with disincentives for the establishment of an evaluation culture.

As a form of situational analysis to inform the path forward, EERE's evaluation lead documented a list of disincentives from program staff across all EERE offices. The findings from the situational analysis included a long-standing focus by management on outputs rather than outcomes; poor quality results information that were not necessarily questioned; little to no interest in learning and adapting; the likelihood of penalization for programs reporting unfavorable results as opposed to not reporting any results; misalignment of program activities due to unclear underlying theory of change; frequent staff turnover from programs; and budgetary decisions misaligned, and sometimes actively adverse to, programs that had taken the initiative to conduct evaluations.

A multi-pronged approach

Figure 1 presents the framework developed in response to the situational analysis. It was developed as a long-term, dynamic framework, amenable to modifications to meet changing needs within the organization. Its underlying vision is aligned (through multiple iterations of logic models and subsequent metrics development) with the overarching goals of the organization, and with best practices culled from U.S. federal agencies.

It aims to facilitate a culture where evaluations provide data for learning and accountability; where data collection is routinized to reduce cost, promote collaboration with stakeholder partners and provide consistency; where impact evaluation methods are consistent and meet high standards of rigor and quality; where accountability functions are centralized and accessible; where staff roles and responsibilities are clear; and above all, where effective use of evaluative information permeates the organization.

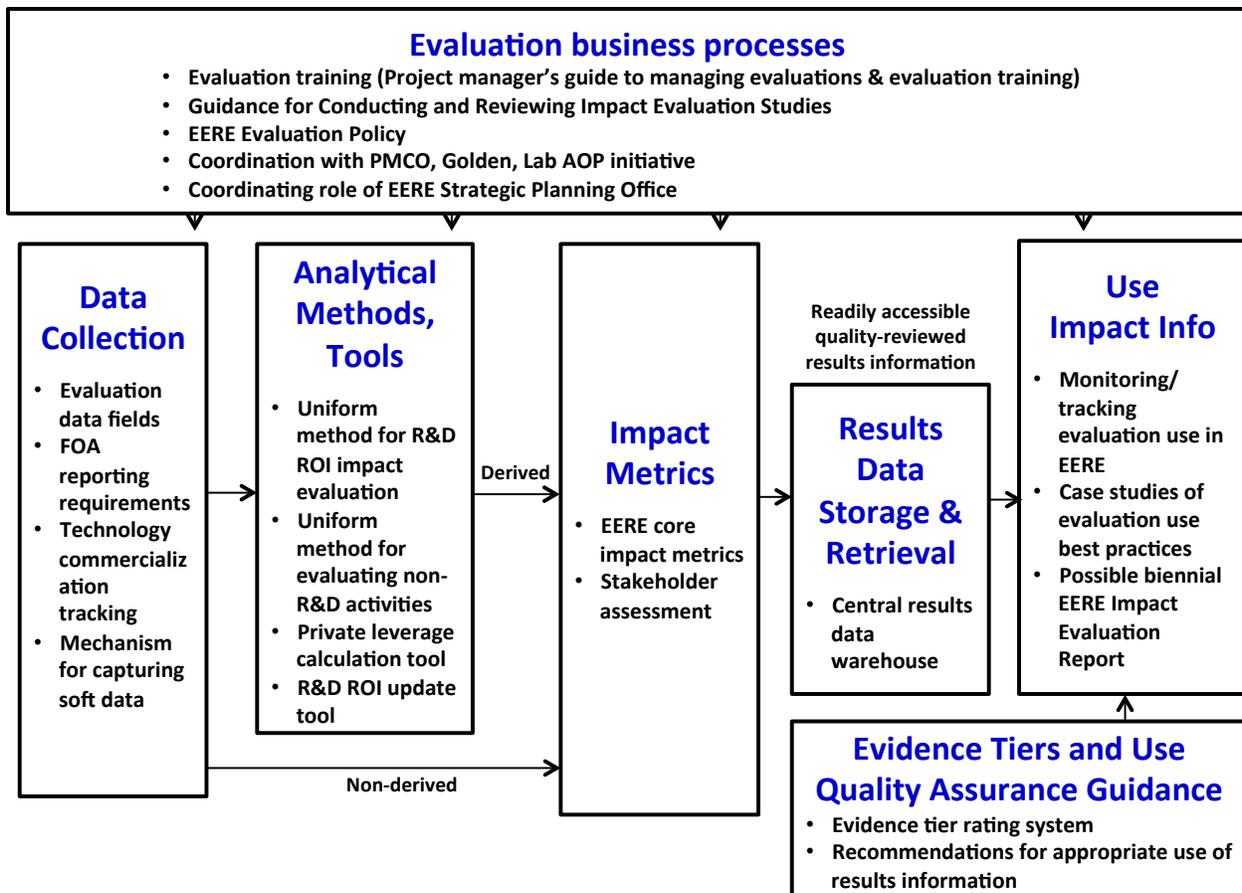


Figure 1. EERE Systemic Evaluation Approach

The actual approach has been to embrace an incremental, multi-pronged strategy, focused simultaneously on the development of standards for evaluation practice, capacity-building in evaluation practices for organizational staff, and commissioning independent evaluation studies as a way of showing by doing. Mayne (2008) citing Kim (2002), notes that every organization has an existing, inherent culture, which reflects, among other things, “what is valued, the dominant leadership styles, symbols, the procedures, routines and the definition of success that make an organization unique.”

Building an evaluative culture involves, for a start, understanding the nature of the existing organizational culture, its history and systems of rewards and sanctions (a form of situation analysis) and tailoring an approach that is responsive and appropriate to the context of the organization. It also involves the vital elements of leadership, organizational support structures, and a learning focus, as described by Mayne (2008).

The experience from EERE lends support to this observation. We began our incremental implementation with a careful targeting of sympathetic program leads, with whom we partnered to co-sponsor evaluations of some of their investments. When this proved sorely inadequate, we made use of the financial crisis of 2007-2008, and the subsequent budgetary cutbacks to gain the interest of upper management within the organization to push for more evaluations. We developed messages to leadership to sell evaluation as the useful tool for efficient and effective management that it is, and emphasized the results from completed studies.

Although we acknowledged the inherent risk of using a crisis to forward an evaluation culture, the crisis had presented an opportunity that could not be allowed to go to waste except through callous irresponsibility. We moved to develop or update accessible guides and standards to support the conduct of evaluation, drafted an evaluation policy for the organization, and commissioned the development of a comprehensive set of metrics across the different programs, including a list of “core metrics” with applicable to the core mission of the organization. We began a series of evaluation trainings for program staff, thus meeting an initial demand for core capability development. Before long, one key use of evaluation results – for accountability – was realized, as upper management began taking data from our commissioned studies to their Congressional testimonies.

Current achievements

We offer a brief compendium of achievements in the period since the initial situational analysis was undertaken, over a decade and a half ago. Since then, EERE has commissioned studies that have resulted in the independent evaluation, by third-party evaluators, of about 30% of the total EERE R&D investments over the period spanning 1976-2013. Some of the results from these independent assessments have been spectacular. For example an impact evaluation of the Vehicle Technology Office’s investment in combustion engine R&D found \$69 billion net benefits (from 1986 to 2007), achieved from the fuel savings to users of heavy-duty diesel trucks. The savings came from the increased efficiency of vehicles due to the laser and optical diagnostic technologies and combustion modeling sponsored by the program. The benefit-to-cost ratio was 53 to 1, based on the approximately 17.6 billion gallons of fuel saved through those investments, amounting to roughly 177 million metric tons (MMT) of avoided CO₂ (Link, 2010).

Sixty-six evaluations (R&D & non-R&D) have been commissioned since 2004. The evaluations peaked in the 2008-2011 period, due to a combination of EERE corporate new evaluation starts (base appropriations) and several new ARRA evaluation studies. The rigor of EERE evaluation studies have improved, and more programs haven began committing funds to evaluate aspects of their investments. Excluding peer reviews and ARRA-funded evaluations, EERE spent roughly \$1.15M annually for impact and process evaluation over the past decade. While this represents a mere .06% of total EERE budget to determine impacts and identify opportunities for making continuous program improvements, it is a step in the right direction for an organization that otherwise had existed for nearly 40 years with little in way of rigorous, independent, evaluation studies.

An evaluation policy document has been developed for the organization, along with several guides and standards, including notably the “Peer Review Guide,” the “Project Manager’s Guide for Managing Impact and Process Evaluation Studies,” and “Evaluating the Realized Impacts of DOE/EERE R&D Programs: Standard for EERE Impact Evaluations.” Additional key guides and standards include “Evaluating the Realized Impacts of Non-R&D Programs: Standard for EERE Impact Evaluations,” the “Framework for Evaluating R&D Impacts and Supply Chain Dynamics Early in a Product Life Cycle,” and the guide for “Rating the Level of Rigor of EERE Evaluation Studies.”¹

¹ All EERE guides are made available upon final review, and can be accessed at: http://www1.eere.energy.gov/analysis/pe_evaluation_documents.html

Lessons learned and a path forward

It is undeniably true that the evaluation lead in EERE has incrementally moved the organization towards an improved appreciation of, and demand for, evaluation information for learning and accountability. Notwithstanding, there is a long road to travel to make a culture of evaluation a realized vision within the organization. We find the lessons learned thus far in the ongoing effort to instate an evaluation culture in the organization to be almost perfectly aligned with the measures prescribed by Mayne (2007) as necessary for building an evaluative culture, and briefly discuss them.

Leadership

Leadership is, without any doubt, the most important factor in the quest to establish an organizational culture of evaluation. The EERE experience has been a case study of how leadership can easily move an evaluation agenda forward, or stifle it. For all the years of endeavor by the evaluation lead, the greatest leaps forward have occurred when senior management have lent their support and commitment, demanded valid results information and made the need for evaluative information a priority. In the particular, sensitive political context of federal agencies, no factor is more critical.

Whether evaluation staff can actually do anything to motivate organization leaders is less clear, because in the context of federal agencies with multiple mandates and shifting priorities, context often determines what is possible. Most political appointees arrive into their leadership roles with mandates and a vision for what they aspire to achieve in the relatively brief period of their appointment. If their vision does not involve a prioritization of an evaluative culture, or worse, a hostile attitude towards it, there is little that evaluation champions can do. Perhaps, however, in that conundrum lies the way: evaluation champions should stay diligent and prepared, pursuing foundational work in less favorable times, while staying ready to seize any opportunity that might present in more favorable circumstances.

Organizational support structures

It is this particular need for supportive organizational systems, practices and procedures, which should motivate the drive to stay committed and diligent, even during periods when leadership is not as engaged. The structures need to be in place when the opportunities arise. An ongoing situational analysis, coupled with creative strategic approaches for addressing observed needs, must guide the development supportive structures. And these structures must be aligned to the goals of the organization to provide some consistency for the culture being developed. Program managers, as an example, must not be shackled to manage only for outputs, but given the autonomy to manage for results, without sanctions. Evaluation information that can be gathered routinely must be included in any established avenues for routine data collection, and evaluative results must be valued as learning tools indispensable to daily management, not as occasional activities done to meet an external requirement. The EERE evaluation lead found great benefit from engaging program leads over long periods, sharing the benefits of evaluation with them, gradually persuading them about the need to evaluate their investments, and eventually providing sponsorship support for the evaluations. In time those program leads began setting aside funds for evaluation.

A focus on learning

Institutionalized learning events are advocated as one key element of what it means to

cultivate a learning focus. This is one of the prongs of the framework for developing evaluative capacity as part of the cultivation of an evaluative culture. The first set of evaluation training offered in the organization was well received, and resulted in staff demand for additional classes. Knowledge sharing emerges as another key development in the strategy, with some program staff having taken the lead in commissioning and managing evaluation studies from their own office and serving as additional resources to other staff within the organization. Related to this, the evaluation lead has assumed more of an advisory role in order to give individual program staff the opportunity to lead the management of commissioned evaluation studies, thus encouraging learning through experience.

Establishing a culture of evaluation is arduous and tasking. It is the rare organization that accepts the responsibility to transparently allow unbiased assessments of its performance, and revels in the advantages facilitated by a culture of evaluation. Nevertheless, a considered, patient approach to instating a culture of evaluation can yield considerable results. It must be guided by a situational analysis, even an informal one, for understanding the scope and scale of the need, the forms and weights of obstacles, and the potential opportunities for action. The subsequent strategy must be bedded in a principle of self-reflection, evidence-based learning, experimentation and willingness to change. It must actively seek senior management support, tirelessly propound and reinforce the merits of results-based planning and reporting, develop the capabilities of staff, and provide positive support for results-based management.

Acknowledgement

Jeff Dowd, Evaluation Lead in the U.S. Department of Energy's (DOE) Office of Strategic Programs in the Energy Efficiency and Renewable Energy (EERE), has been the steady hand at the wheel for cultivating a culture of evaluation in EERE. This paper is an ode to his tireless effort. In this, he is an exemplar public servant.

Reference

- Campbell, Donald T. 1991. "Methods for the experimenting society." *American Journal of Evaluation*, Vol. 12 (3): 223-260.
- Kim, P. S. 2002. *Cultural Change in Government: Promoting a High-Performance Culture A Review of Ten Years of Modernisation: The HRM Perspective*. Human Resources Management (HRM) Working Party Meeting Paris: OECD.
[http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=PUMA/HRM\(2002\)11&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=PUMA/HRM(2002)11&docLanguage=En)
- Link, A. N. 2010. "Retrospective benefit-cost evaluation of U.S. DOE vehicle combustion engine R&D investments: Impacts of a cluster of energy technologies." Washington, DC: U.S. Department of Energy Office of Energy Efficiency and Renewable Energy.
- Macadam R, Drinan J, Inall N and McKenzie B. 2004. *Growing the capital of rural Australia: The task of capacity building*. RIRDC Publication No. 04/034. Canberra.

- Mayne, J. 2008. *Building an evaluative culture for effective evaluation and results management*. ILAC Working Paper 8, Rome, Institutional Learning and Change Initiative. www.cgiar-ilac.org.
- Mayne, J. 2007. "Evaluation for accountability: Reality or myth?" In *Making Accountability Work: Dilemmas for Evaluation and for Audit*. M. L Bemelmans-Videc, J. Lonsdale and B. Perrin, Eds. New Brunswick, NJ: Transaction Publishers.
- OECD. 2006. *The challenge of capacity development: Working towards good practice*, OECD DAC Network on Governance.
- Simister, N. and Smith, R. 2010. *Monitoring and evaluating capacity building: Is it really that difficult?* Praxis Paper 23. Oxford: International NGO Training and Research Center.
- Trochim, W. M. K. and Donnelly, J. P. 2006. "An evaluation culture." *The Web Center for Social Research Methods*. <http://www.socialresearchmethods.net/kb/evalcult.php>