Evolution of EM&V: Moving Towards a Systems Design

Brandy Brown, CLEAResult

ABSTRACT

This paper envisions energy efficiency programs buttressed by systems design-focused evaluation and presents an argument for evolving current EM&V practice. The changing energy landscape abounds with new technology, behavioral considerations, and a flood of big data. Although traditional energy efficiency evaluation practices are well suited for verifying program achievement and serving as a regulatory check, they are less suited for developmentally improving programs. Evaluation practice needs to evolve to address both the external regulatory accountability needs of public service commissions and the internal program improvement needs of energy efficiency programs. Building on traditional evaluation, monitoring, and verification (EM&V) practices, this paper advocates for the expansion of evaluation practice to include internal evaluation thereby opening the door to internal systems that facilitates learning, adaptation, and growth of programs. Expanding the evaluation practice to include internal evaluation can help facilitate real-time comparisons of indicators such as satisfaction, awareness, and energy savings would provide a more responsive form of evaluation that utilities seek. Integrating evaluation systems into energy efficiency programs is the next step in the evaluation of energy efficiency evaluation practice and excellent way to reap the positive benefits of this new era of *Big Data*.

After explaining how such a system facilitates rapid, responsive evaluation, this paper presents a short case study of a current energy efficiency monitoring and evaluation system. Additionally, it provides lessons learned in designing, implementing, and maintaining an advance monitoring and evaluation system to help determine the effectiveness and value of each energy efficiency program.

Introduction

Achieving high-energy savings is not just about new technology and big data; it's also about creating evaluation systems that mold information into actionable intelligence. As we proceed in the era of *Big Data*, evaluation must evolve to meet the challenge and reap the rewards. First, evaluators must recognize and tackle the contradiction or ambiguousness of our purpose: *Is the purpose of evaluation to provide accountability or is it to approve programs?* On one hand, evaluation serves a regulatory accountability function to ensure that programs are meeting statutory requirements. One the other hand, utilities and program administrators would like to use evaluation to inform continuous program improvement. Program managers are clamoring for evaluation to serve a developmental function and help programs improve. However, evaluation that serves a developmental function can differ greatly from evaluation that serves a regulatory function. Where evaluation serving a regulatory function needs to maintain the appearance of objectivity, evaluation serving a developmental function is less concerned with appearing objective and instead focuses on painting the most accurate picture in order to inform strategic change. Also, evaluation serving a regulatory accountability function needs to occur externally to the program, whereas evaluation serving a developmental function can occur within

the program and by internal evaluators. *How can we meet divergent needs and challenges of efficiency programs and simultaneously reap the rewards?* To better recognize and respond to the changing energy landscape, we need to progress rather than conduct business as usual (Friedmann, 2011). Continuing to rely on post-hoc data collection will impede our awareness of energy savings issues and thus, our response to them. In order to more responsively adapt to future challenges in achieving energy savings, we need continuous, real-time snapshots of key performance indicators such as energy savings, costs, satisfaction, and awareness. Developing systems to generate instantaneous program data would advance energy efficiency evaluation enabling it to provide timelier, actionable feedback to energy efficiency programs. The practice of energy efficiency evaluation as we know it should expand to serve the competing needs of energy efficiency programs by acknowledging the distinction between internal and external evaluation– external accountability focused evaluation and internal program improvement focused evaluation. Although the subject of this paper is the evolution of energy efficiency evaluation practice, it is focused on expanding evaluation within in programs rather than modifying evaluation practice within the third-party evaluation realm.

As programs adjust to new technology and data streams, it becomes more critical to incorporate evaluation within program design. Specifically, evaluation must be structured as a system that comprehensively and responsively addresses the unique priorities of each energy efficiency program (Patton, 2011). This internal evaluation internal evaluation can take many forms, however, this paper envisions it as a monitoring and evaluation system.

A monitoring and evaluation (M&E) approach focuses on strengthening program operations by building organizational capacity in data, analysis, and evaluation (Stufflebeam & Coryn, 2012). A monitoring and evaluation system is central to undertaking evidenced-based management, program design, and budget decisions. Specifically, M&E is designed to assess performance on specific goals and enhance evaluation use by incorporating data collection activities into program design (Mathison, 2005). It can provide accurate, real-time program snapshots on program performance and auxiliary indicators such as customer satisfaction. M&E systems also provide a sustainable way for programs to track and adjust program outcomes. As programs operate, program data is captured, internally evaluated, and disseminated to management. This arrangement facilitates a closer working relationship between program implementation and evaluation, thereby reducing the burden of external evaluation on the program staff and maximizing the usefulness of evaluation.

Monitoring is the ongoing process of collecting data on program performance toward goals and objectives as a program progresses (Mathison, 2005). Program managers often review program scorecards progress on an incremental basis, but this is not monitoring – it is reviewing. Monitoring involves strategic management in response to real time data. In other words, monitoring is concerned with; "*Are we taking the proper actions?*" not only "*Are we making progress on savings goals?*" In this comprehensive approach, monitoring is an on-going process that focuses on the relevance of program metrics to savings goals, the effectiveness of energy efficiency programs, contribution and worth of programs or measures to the overall goals, key drivers that contribute to success, aiding intra-organization coordination to reduce transaction costs, risk mitigation, and sustainability of results. The lessons from monitoring are discussed periodically and used to inform actions and decisions.

While monitoring provides real-time information desired by program managers, evaluation provide more in-depth assessment (Scriven, 1980). Evaluation answers the questions that arise from monitoring data. For example, if a program manager notices unexpected

variations in program performance, she can use the data generated through Monitoring to investigate potential reasons for the variations. The time-series data collected through monitoring can help her to determine the impact of contextual factors on program performance (Shadish et al., 2002). Unlike third party evaluation, the "E" in M&E denotes internal evaluation; evaluation conducted within and by the organization implementing energy efficiency programs.

The ultimate objective of an effective M&E system is to support accountability to clients and regulators, promote risk management, facilitate prompt corrective action, and enhance organizational learning (UN, 2009). Simply put, monitoring and evaluation systems are internal implementation tools used to provide consistent data on what is and isn't working. Programs with strong monitoring and evaluation systems tend to stay on track because problems are often identified earlier. Early problem identification reduces the likelihood of having large variances between gross and net savings later on.

Monitoring and evaluation can play a major role in enhancing the effectiveness of energy efficiency programs by helping relate program progress to program plans. Proper planning should include well-defined program goals to elucidate which indicators to track and how to conduct monitoring. In the absence of careful and accurate monitoring of program data, evaluation can't be done well. Monitoring facilitates evaluation, while evaluation uses additional insight and data to apply a framework for analysis. Expanding the existing concept of evaluation practice to include internal evaluation would facilitate real-time comparisons of indicators such as satisfaction, awareness, and energy savings would provide a more responsive form of evaluation that utilities seek. This expanded system would also provide timely and accessible systematic reviews of energy efficiency programs.

Envisioning Systems-Design Evaluation

The next generation of EM&V will be marked by well-designed, responsive evaluation systems that provide sustainable solutions for tracking and modifying program performance; incorporating monitoring and evaluation is just one aspect of evaluation practice becoming more responsive to program needs. However, incorporating monitoring and evaluation practices is more easily said than done.

The first shift our industry would need in order to incorporate M&E involves the role of evaluation in energy efficiency. In order to achieve responsive evaluation, our foundation of EM&V work needs to shift from being reactionary to a symbiotic relationship with implementation. Establishing a partnership between evaluation and implementation will allow evaluators to work with programs from their inceptions and provide implementers the real-time evaluative insights they seek. By beginning with a partnership between evaluation and implementation of energy efficiency programs, we can begin to develop advanced monitoring and evaluation systems to help organize and extract relevant information from past and current program activities. This data can then be used to re-calibrate programs and orient future program designs. This system allows us to judge if programs are preforming as expected, allow us to shift a program's direction based upon external factors, understand what is and isn't working with our programs (before external evaluation), and direct us on how future efforts can be improved. To achieve responsive evaluation we need to implement M&E systems.

The next hurdle in moving towards more responsive evaluation is shifting the how we perceive the types of evaluation. For example, efficiency evaluation typically takes the form of process evaluations or impact evaluation by which evaluators retroactively "measure" program effectiveness through energy savings and other stipulated indicators. But what if we applied

evaluation methodology that allows us to be responsive to program needs? What if we could monitor and evaluate programs in a real time basis? How would program managers use these learnings to improve their implementation practices? These questions not only push us towards Developmental evaluation, it also pushes us to examine the assumed sequential nature of program operations followed by evaluation. Developmental evaluation is an evaluation approach that aims to responds to dynamic environments. It is carried out by an internal evaluation team that is integrated into the program process. Implementing M&E system would help us break the sequential cycle of evaluation-after-implementation by allowing us to evaluate during implementation.

What Does a Monitoring and Evaluation System Look Like?

An M&E system can take many forms depending on the structure of the programs of focus. It can be dispersed through an organization or centrally located in an internal evaluation unit. For the purposes of this exercise, envision the product of an M&E system as periodic report cards. The system that generates this report card consists of internal and external data collection, analysis, and evaluation tasks. For example, imagine a system of organizational policies and procedures that integrate human understanding and technical capabilities to facilitate rapid evaluation of program performance. This system, which integrates transactional data collection, information management, and routine analysis of program progress toward goals to allow the program to reflect on results as the program is operating. This is a monitoring and evaluation system.

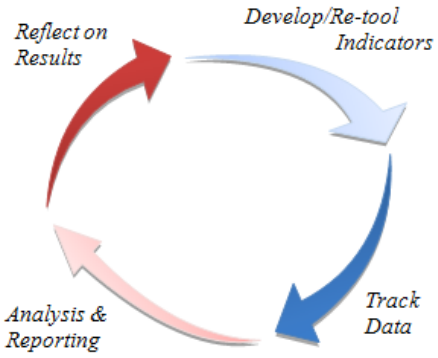

Figure 1. Conceptual M&E system.

Partnership with External Evaluators

Although the M&E goals are driven internally by implementation staff, there is still a function for external evaluators to play within an M&E system. Implementation staff would be remiss to only share the information internally. For implementers, M&E is a management tool that allows *program staff* decision makers to make evidence based management decisions during project implementation and planning of future programs. Additionally, monitoring is a means of internal accountability towards interested stakeholders.

One of those stakeholders is the external evaluator. Monitoring and evaluation can provide external evaluation teams with highly detailed insight into program operations. This insight includes information about the programs in question. Indicators include customer satisfaction, customer awareness, program cycle time, and energy savings, This M&E system is used to integrate the different perspectives of implementers and evaluators to move closer toward program success.

External evaluators can play a substantive role in the M&E process by validating the quality and reliability of the implementer's M&E data. There is the impression that the implementer would collect data that is more biased in favor of program success as compared to

data collected by an external consultant. Therefore, it would be to the benefit of implementers to engage external evaluators at the start to validate the internal data collection methods to ensure that internally, the program is compiling an accurate picture of itself.

A Case for Evolution

This section describes an existing monitoring and evaluation system and the lessons learned designing, implementing, and maintaining the system to facilitate responsive evaluation. It is based on a Company that has implemented an M&E system. The discussion that follows applies to any energy efficiency portfolio.

Project Background

Company X is piloting an advanced M&E data system that allows them to monitor and evaluate program performance at the most granular levels. This structured data system is the foundation for monitoring and evaluation. It empowers responsive management and internal evaluation of energy efficiency programs.

The project is currently in its pilot phase and has three objectives:

1. Creation of a real-time performance tracking system
2. Empower program managers with information to adjust programs mid-course
3. Promote a culture of continuous improvement through organizational development, by building the capacities of program implementation staff.

Development

The M&E system developed by Company X examines program portfolios and has five central components: a cost benefit analysis of all projects at the measure level, custom performance indicators, comprehensive management reports, internal evaluations of programs, and comprehensive spending reviews. Each component is tracked using a relational database that allows data to be broken down and analyzed based on a custom set of indicators. Using a relational database is important because it allows users to avoid data duplication and is more amenable to future updates that require record expansion or editing.

The internal evaluation team worked to engage the program managers during the M&E planning phase. Special consideration was given to explaining that the purpose of M&E is to aid the continuous improvement of the programs and adaptive learning of the implementation team. The internal evaluation and program staff collaborated to develop which performance indicators were most relevant for each program and prioritized those indicators. These metrics were also reviewed with external evaluators to ensure that they were properly measured and recorded.

The M&E indicators were defined during the project baseline and tweaked to align with prior year evaluation results. Data routinely collected by the programs are validated, analyzed and reported in the M&E database.

Management

The M&E system is internally implemented and continuously managed by the implementation team. However, the outputs of the system are shared with external stakeholders including third party evaluators. The system is financed solely through the implementation budget as part of the internal quality assurance functions.

All M&E functions are tracked in a company-wide relational data base at the customer and measure level. This allows the company to review at a national level program performance by program type and region. It also allows for a meta-analysis of performance results nation-wide.

This M&E system enables real-time, responsive evaluation of energy efficiency program activities by simultaneously tracking various indicators at each customer touch point. Managers can query the system by different variable combinations (i.e. client, program type, region, fuel type, etc.) to analyze and report program performance data. The M&E system allows users to monitor and evaluate output and impact indicators of projects against baselines and stated goals. Additionally, indicators of risks to savings and process efficiencies are tracked to allow for timely corrective action. High level management is empowered with the ability to make evidence-based decisions about impact, efficiency, and effectiveness of their programs during the program cycle. This advanced M&E system is user-friendly and customizable for each client. System features include online data collection, streamlined analysis, custom reporting, and transferability to statistical software packages to facilitate further analysis.

Benefits

In addition to providing program staff with real-time insight into program operations, this internal M&E system also benefits external stakeholders. For example, the external evaluation team benefited from easily accessible and well-organized program data. Rather than waiting for program staff to gather and share program data, the external evaluation team can examine program operations at their convenience. Data requests are no longer cumbersome processes of multiple meetings to resolve issues with incomplete data. Since each record in the M&E system's relational database is attached to all the supporting program documentation, external evaluators can more easily sift through program data to ascertain which program components need improvement.

Additionally, policy makers and regulators benefit from the most recent information on program performance. Rather than using evaluation results from past program years, these stakeholders are able to make policy judgments with a more real-time picture of program operations. In fact, they get access to pilot program performance months after pilots are launched as opposed to years after launch. This allows for more real-time adjustments to enable programs to respond sooner to shifting market conditions.

Lessons Learned

At this stage of the M&E system implementation there are no definitive conclusions that can be presented. However, there are considerations and learnings from this case. Implementing an M&E system can be particularly challenging. For example, the initial planning phase requires a large investment of human and technological capital from which the value may not be immediately realized. Additionally, the shift required in the organizational culture of

implementers and evaluator may require a fair amount of training including an introduction to the concepts of M&E as well as experiencing the benefits in order to understand the value.

The first months of developing an M&E system should focus on capacity building in:

a) Participatory tools for creating indicators to monitor the program and whether goals have been achieved or not.
b) Reporting skills that adequately communicate performance, outputs, and outcomes.
c) Planning tools that incorporate evaluation findings in planning and implementing future programs.

These lessons in developing and utilizing real-time monitoring were also exhibited in the launch of Efficiency Vermont. In 2000 Efficiency Vermont launched its first year of a multi-year contract by developing and implementing real-time monitoring of energy efficiency adoption in over 30 utility territories, across various customer classes (Hamilton et al, 2002). In this performance monitoring system, Efficiency Vermont tracked 35 measures with specific targets. The program performance indicators focused on program results, market effects, and activity milestones and were tied to the contractor's performance incentive.

Both the Efficiency Vermont & case study show how monitoring and evaluation aid traditional evaluation approaches by establishing systematic approaches to gathering, measures, and understanding how energy efficiency programs work..

How to Develop a Monitoring and Evaluation System

Developing an M&E system, much like the system itself, is an iterative, ongoing process. As such, the steps in developing a monitoring and evaluation system are also iterative, and continuous. The following steps describe the general development of an M&E system:

1. Determine Resources
   a. This step helps an organization understand what current financial, human, and technological resources could potentially support the development and implementation of a M&E system
2. Identify existing organizational M&E functions
   a. This step allows the organization to understand the current strengths and weaknesses. It also allows organizations to adjust current processes to yield a clearer picture of real-time program performance. It will also help clarify what information is needed for what purposes, and when this information is needed.
3. Define performance indicators – Performance indicators serve as the measurable gauges of whether a program is meeting established criteria. Therefore, indicators serve as the underlying basis to determine whether the program is successful. Potential indicators include savings, cost-effectiveness, customer satisfaction, program cycle time, and participant retention.
4. Develop a monitoring matrix that defines what you are monitoring, the data source, how information will be obtained, schedule and frequency of monitoring, level of detail, responsibilities, means of verification, resources required for this task, and risks and assumptions associated with the monitoring activity.

5. Develop a method to track indicators that is secure and consistent. Specifically, decide where will the information be stored, how frequently will it be updated?
   It is important that progress on key indicators is kept up to date and accurate. As such, it may be necessary to assign this task to one individual. The method by which indicators are track may include an organizational policy on frequency of indicator updates, roles and responsibilities of team members, and the format or software package that is used for tracking.
6. Assign ownership of the M&E functions
   a. Data entry - Data entry is conducted at customer touch points. Data on program performance indicators is inputted into the relational database. This data base allows users to pull, on a real-time basis, data that speaks to specific indicators on the measures and customer level.
   b. Review - Program managers can review program performance at any time and take corrective action for any issues that arise through existing Quality Assurance and control protocols.
7. Conduct periodic reviews to ensure that all necessary data is collected and supplied in the most appropriate formats.

Does Your Organization Have the Capacity for an M&E System?

The capacity for monitoring and evaluation, like technical specialty areas, depend on a fertile environment, organizational readiness, and individual capabilities. The need for multilevel capability within an organization requires a holistic perspective of whether an organization is ready to coordinate monitoring and evaluation.

The following questions may help you identify whether your organization is ready for an M&E system:

1. Does high-level management support evidence based decision making?
2. Are sufficient resources available for monitoring and evaluation including staff with statistical and analytical skills?
3. Is program information at the customer and measure level available for analysis?
4. Can information be disaggregated into indicators of performance?

If you answer yes to all of the above questions, your organization may be ready to implement an M&E system. If, your organization is not ready to implement an M&E system, there may be areas where you can incorporate singular monitoring and evaluation checks (single data collection points) that provide feedback on program performance.

Conclusion

The changing energy landscape requires evolution; it requires continuous, real-time snapshots of program and quick response to mitigate risk. This evolution of energy efficiency evaluation is central to the future adaptation and growth of the industry as a whole. Actionable intelligence can advance future achievement of energy savings, but first we must build evaluation capacity. This includes cementing a partnership between evaluation and implementation and employing smart data systems. By expanding current evaluation practice to encompass internal evaluation systems, evaluation can help programs become more successful

and cost-effective. Systematizing internal evaluation would also reduce many evaluation-related burdens placed on program staff and increase their availability for other tasks that advance program success. Monitoring and evaluation is just one piece of this future.

Monitoring and evaluation is a tool that can facilitate real-time comparisons of common factors such as satisfaction, awareness, and energy savings to provide a more responsive form of evaluation and reduce many evaluation-related burdens. It is not a replacement for third party EM&V; rather, it is a compliment to it. M&E users must be cautious in drawing conclusions based on glimpses of this continuous data stream as lags between interventions and the programs ability to measure achieved savings results can skew real time data. To avoid pulling the plug on programs prematurely, program managers must take the programs characteristics and context into consideration. A hybrid approach, that pairs an internal M&E system and external evaluation, would help produce information that implementers need and can use to improve programs closer to real-time

The positive implications associated with broadening evaluation to embrace M&E systems include, faster feedback, empowered evaluation, tracking customers through the program cycle, and integrating evaluation into each customer touch. Additionally, this type of system would allow us to respond to the growing demands for client and regulatory accountability, reduce inefficiencies, and encourage programs to focus on sustainability of specific measures within a portfolio. A systems design approach to energy efficiency evaluation is a logical step in the evolution of our energy efficiency evaluation.

References

Friedmann, R. (2011). A Fresh Look at Evaluation to Support Energy Efficiency in the 21st Century. *IEPEC.*

Hamilton, B. Plankuett, J. Wickenden, M. (2002) Gaugin Success of the Nation's First Efficeincy Utility: Efficiency Vermont's First Two Years. Asilomar, CA. ACEEE

Mathison, S. (2005). *Encyclopedia of Evaluation.* Thousand Oaks, CA: Sage.

Patton, M. Q. (2011). Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use. New York, NY. The Guiliford Press.

Scriven, M. (1980). *The Logic of Evaluation.* Inverness, CA: Edgepress.

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

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston, MA: Houghton Mifflin.

Stufflebeam, D. L., & Coryn, C. L. (2012). *Evaluation, Theory, Models, & Applications* (2nd ed.). San Francisco, CA: Jossey-Bass.

United Nations Development Program. (2009) *Handbook on Planning, Monitoring, and Evaluating for Development Results*. New York, NY. United Nations.