Shapeshifting Evaluation: Rapid, Flexible and Actionable Evaluation Within a Context of Continuous Improvement

Hannah Justus, EMI Consulting, Seattle WA
Kerry Meade, EMI Consulting, Seattle, WA
Erik Lyon, EMI Consulting, Seattle, WA
Mahdi Jawad, Consolidated Edison, New York, NY
Rosanna Jimenez, Consolidated Edison, New York, NY

ABSTRACT

In the face of ever-increasing energy efficiency and demand reduction goals, utilities are adapting by developing continuous improvement programs (CIPs). By making their programs more agile, utilities can react in real-time to improve programs and meet goals, but that can leave process evaluations one step behind. Evaluators can find themselves providing feedback on last year's program design when the current program design has already changed. CIPs can also result in misalignment with regulatory reporting requirements seeking summative process evaluations that provide snapshots of a program, not a moving picture. All of this can lead to frustration and confusion between program implementers, evaluators, and regulators.

Introduction

Understanding how to successfully conduct evaluations within a shapeshifting program context will be critical for the next generation of energy efficiency evaluations–and equally important for utilities to achieve their energy efficiency and demand reduction goals. This paper discusses lessons learned for mitigating challenges that emerged during the process evaluation of a CIP at Consolidated Edison (Con Edison), the Demand Management Program (DMP). The program design and processes rapidly changed throughout the course of the evaluation to meet program goals more efficiently. This nimble implementation strategy allowed the program staff to do everything they could to meet goals, but it posed challenges to the evaluation with regard to communication, data collection, and reporting. This paper provides recommended approaches to managing evaluations for CIPs, including communication techniques, research methods, and reporting styles.

The demand-side energy management landscape is rapidly changing due to a number of factors including policy changes, decommissioning of legacy coal and nuclear plants, new technologies, and increased energy efficiency goals. At the same time, utilities are modernizing to offer customer-centric services and use data to drive decision making. This has led to more utilities incorporating continuous improvement protocols into their demand-side management program designs. By making their programs more agile, utilities can react in real-time to improve programs and meet goals; but given the traditional "snapshot in time" approach to program evaluation, this can leave process evaluators one step behind.

Historically, rigorous process evaluation relied on stable implementation environments, whereby evaluators can use a sample of program participants to collect program data that remains relevant when it is later analyzed and results are presented. When program designs are demanding more rigorous process evaluation to support critical system management needs, program evaluators are challenged to provide program implementers with evaluation results that are timely and informative while still meeting highly rigorous evaluation standards and oftentimes summative regulatory reporting requirements.

This paper provides recommended approaches to managing process evaluations within a continuous improvement program (CIP) context. It highlights methods the authors relied upon, as well as lessons learned within the context of the Consolidated Edison (Con Edison) Demand Management Program (DMP) process evaluation. During the evaluation, the authors identified three evaluation elements upon which the success of a CIP evaluation hinged: the need for active and regular communication between evaluators and program staff, unique data collection and sampling methods, and varied reporting methods. This paper first presents the context of the Con Edison DMP. It then explores each of the three elements and concludes by presenting key lessons and recommendations for future CIP process evaluations.

Program Background

With a need to address peak demand needs, utilities are turning to energy efficiency, and more specifically demand-focused energy efficiency. While energy efficiency is a valuable resource for managing system load, using it as such necessitates greater scrutiny and careful management to ensure the intended program goals are achieved. Such a novel program is the model candidate for a CIP. Con Edison, the primary utility provider for New York City, developed the DMP to determine whether it could leverage peak-demand focused energy efficiency to address a projected supply shortfall if a nuclear facility, the Indian Point Energy Center, was decommissioned. The DMP is a unique systems-based solution that targets coincident peak demand during the summer months. The program was distinctive in that it had a one-time target of achieving 100 MW of installed permanent summer peak demand reduction by June 2016, and would then terminate as a program. To meet this goal, the DMP provided financial incentives to Con Edison commercial electric customers, and third-party developers acting on the behalf of the Con Edison customer, to implement qualifying energy improvements within their facilities.

Since Con Edison had limited time to meet their goal of reducing demand by 100 MW, the DMP required nimble and proactive attention to customer experiences so program staff could quickly adjust processes and offerings to best meet their goal on time. At the same time, the utility needed a rigorous process evaluation that would sustain regulatory scrutiny and provide rapid feedback to inform continuous improvement opportunities. This type of process evaluation is easier to perform when a program is static, however the DMP was necessarily evolving. As a result, the evaluation team developed several protocols to provide real time feedback and adaptive evaluation approaches. The unique elements of the DMP evaluation that required these approaches are explored in the remainder of this section.

Unique Elements of Continuous Improvement Programs (CIP)

For a utility to operate a CIP, the utility must design its program in a way that supports continuous improvements to the program design and/or program implementation. This section describes three unique elements of the DMP that the authors encountered in the DMP process evaluation, and which the authors believe are common to any CIP: the need for rapid feedback, varying customer experiences, and evolving program documentation. These unique elements are important because they impact how a process evaluation for a CIP can be designed.

Need for Rapid Feedback Evaluation

CIP's by nature require rapid feedback evaluation. Because CIPs are trying to rapidly respond to customer experiences, CIP staff rely on evaluators for rapid feedback to help tune the program as quickly as possible. This is different from a traditional evaluation, where evaluators typically report on findings after data collection is complete, since program adjustments tend to occur in the subsequent year's program iteration. In the CIP context, evaluators need to provide interim results as quickly as possible

because there might not be another opportunity to truly have an impact on program processes—in the case of the DMP, rapid feedback on evaluation activities provided an opportunity for process evaluation results to influence program processes in support of hitting its one-time goal. Rapid feedback evaluation can come in many forms but all emphasis research methods that can be conducted quickly, e.g. interviews with key stakeholders, program data analyses, and best practice research. Equally important is for evaluators to be able to communicate results quickly to program staff. Communication styles can also take many forms but must emphasis providing meaningful results to program staff in a way that does not require significant amounts of time to develop the reporting tool. Rapid feedback communication tools can include memorandums, slide decks, or notes from results discussions with staff.

For the DMP evaluation, the evaluation team provided rapid feedback evaluation in the following ways: staff feedback on processes during project kick-off meetings, process mapping, interviews with participants and trade allies, program data analysis, and common practice research. Each of these techniques provided opportunities for quick feedback to program staff by either engaging staff directly in the evaluation process or by relying on research methods that could be conducted without requiring an extensive amount of planning time.

To engage staff directly in the evaluation process, the evaluation team held a series of in-person meetings with program staff as part of the evaluation kick-off meetings. These preliminary meetings provided an opportunity for staff to immediately identify challenges and process improvement opportunities. Staff were then able to act on these opportunities, allowing them to improve internal program processes straightaway. The evaluation team conducted a second in-person meeting to map out program processes in detail with program staff. The process mapping task allowed a second opportunity for staff to collectively discuss implementation challenges and share process improvement opportunities. Like the preliminary discussions at the kick-off meeting, staff could immediately implement recommendations that came out of the process mapping meeting.

The evaluation team also relied on three research methods that it could conduct without an extensive planning phase. First, it conducted in-depth interviews with five participants and three market actors. The evaluation team targeted these interviews to participants and market actors best positioned to address high-priority research questions. The evaluation team developed topic guides, rather than formal interview guides, to speed planning efforts and allow more opportunity for interviewees to provide most relevant feedback on their experiences. The evaluation team presented results to the staff via a phone call and provided results in a memorandum. It also used interview results to inform the primary data collection tasks. Secondly, the evaluation team conducted an analysis of program data. The analysis informed immediate program tracking recommendations and provided a snapshot into the program's progress towards its reported goals. Third, the evaluation team conducted secondary research on common practices, to better understand how the DMP compares to similar programs. The evaluation team provided all of these research efforts in interim memorandums to the program staff to provide results as quickly as possible. As a result, program staff could implement evaluation recommendations immediately, which helped staff better position themselves to meeting their goal.

Varying Customer Experiences

The nature of a CIP means that the program will change over time–leading to varying customer experiences. This is different from a traditional program where customer experiences are fairly static to the extent that program staff are implementing the program consistently and according to the program design. In a traditional program, evaluators have relative certainty that if they sample a group of participants in one month, results from the analysis will represent overall customer experiences. In the CIP context, however, if an evaluator pulls a sample at one moment, results from that study may look very different when the evaluator applies those findings to the program structure a few months later.

This was evident in the DMP in two ways. First, during the first half of the DMP implementation, contractors estimated energy savings from lighting projects in a variety of ways. Program staff realized that these different approaches ultimately meant that staff needed to spend time with contractors fixing their applications prior to approving the project. This created bottlenecks and frustration among contractors and program staff, and given the one-time program savings goal, this also jeopardized program success. As a result, program staff developed a lighting tool for contractors to more accurately estimate project savings while also using similar approaches and inputs in their estimates. This change led to improved customer, trade ally, and contractor experiences because projects could be processed more efficiently. In this case, evaluators needed to know about this change to help explain impact-related results that they found when assessing the data.

Additionally, Con Edison partnered with NYSERDA to implement the DMP at the beginning of the program implementation but over time, NYSERDA's role became much more limited. After a certain point, Con Edison was responsible for all new project applications. As a result, Con Edison experienced a large project backlog as the project volume quickly increased and had to ramp up their own efforts to independently meet program goals. The DMP hired additional external consultants to support the M&V review process, but it took time to train new staff and work through the backlog of applications that occurred during the brief period. This change was important to recognize because the evaluation team sampled projects around the time of the backlog, which meant that many customers described long waits (sometimes months)- which was true for them, but not for the entire population of program participants.

Evolving Program Documentation

Another unique element of a CIP is the fleeting accuracy of program documentation. Since CIPs are rapidly changing, program documentation can quickly become outdated. While traditional evaluations also sometimes suffer from poor documentation, CIPs are unique in that program documentation, if it exists, can be outdated nearly as quickly as it is finalized. This is an important factor because traditional process evaluations rely heavily on program documentation as a reference to assess the program design, program processes, and program objectives. Evaluators, in the CIP context, on the other hand cannot necessarily rely on program documentation as a resource.

This aspect was highlighted in the DMP evaluation in two ways. First, the DMP allows applicants to aggregate projects together to meet program minimal project size requirements and allow customers to receive bonuses for larger (or multiple) projects. However, the aggregation objectives and processes changed during program implementation. Evaluators needed a better understanding of the changes that took place, and needed to allow significant time to meet with implementers to outline a common language around aggregation processes. Second, DMP implementers originally had a June 2016 deadline for when projects needed to be completed. However mid-way through implementation, program staff began exploring whether this deadline could be extended to December 2016. Like changes to the aggregation process, the evaluation team needed to allocate significant time to clarify this, as it was a critical data point to determining the likelihood of whether enough projects would be completed within the project timeframe to meet their goal.

Managing Evaluations in a CIP Context

As explained above, CIPs present unique evaluation challenges that evaluators need to address for the success of the evaluation. During the DMP process evaluation, the authors identified several strategies to successfully manage the unique challenges of a CIP evaluation. These strategies can be separated into three main categories: communication, research methods, and reporting.

Communication

It is apparent to any experienced project manager that active communication is critical to keeping all project stakeholders adequately informed about key project developments, which typically include evolution of scope, budget, and schedule. But while it may seem obvious, and common practice, to hold a client/evaluator project check-in meeting on a regular basis, knowing who and when to communicate becomes critical when evaluating a CIP.

First, knowing who to include in such a meeting can be the lynchpin to a successful engagement. As energy efficiency program evaluation has become standard practice and a regulatory requirement in many jurisdictions, demands on the program staff have led to many utilities leaving program staff out of evaluation project meetings and even direct email communications with evaluators, meaning the program staff who are privy to the most up to date program implementation details are not included in the project meetings or communications, and likewise, evaluators are not included in relevant program communications. While it is always helpful to ensure evaluators are updated on program changes in a timely manner, when evaluating a CIP, it is helpful, if not critical, to go the extra step to include the evaluation team and program staff in regular check-in calls. Which leads to the second aspect to effective communication: the frequency of meetings between the evaluation team and program staff. Meeting frequency will depend on how fast the CIP is changing- weekly, monthly, or quarterly. Alternatively, these meetings could be held on a more ad hoc basis to account for when the program makes a change to its processes or to account certain stages in the evaluation process. Ultimately meeting frequency can depend on CIP needs, evaluation needs, and staff availability.

Over the course of the DMP process evaluation, the program staff implemented small, as well as large, changes to the program design to respond to program performance, market feedback and in some cases evaluation feedback. Keeping the evaluation team updated on these frequent changes was critical to the success of the evaluation. These types of meetings can help the evaluation team most effectively use evaluation resources and/or identify additional topics for deeper exploration.

Research Methods

During a process evaluation of CIP there are unique nuances to consider when designing research methods. While traditional summative process evaluation is designed to collect data about a relatively stable program over a period, CIPs are, in contrast, constantly changing. Such programs might be different in June than they were in January. To accurately represent the program and customer experiences over this period, it may be necessary to distinguish between data from different times throughout the evaluation period.

This became evident when we looked at data before and after NYSERDA's role in the DMP changed. Contractors who participated before NYSERDA's exit had generally positive experiences with Con Edison and few complaints with the timeframe required to complete projects, while those who participated after the exit frequently indicated dissatisfaction related to delays in the program and difficulty contacting program staff, who were busy working through the backlog. Further, contractors who had participated both before and after often referenced the "early days" of the program before the delays, and in some cases even attributed these delays to NYSERDA's exit from the program. By the time data collection for the evaluation was complete, program staff reported that they had staffed up to handle the backlog of projects, and related issues were now resolved. Altogether, this suggests three distinct phases of the program: during NYSERDA's joint management of the program, immediately after their exit, and after Con Edison's increased staffing levels.

The DMP process evaluation largely reflected the second of the three phases (the first phase was represented to a lesser extend in the rapid feedback research, which was conducted immediately after the evaluation began). The DMP, while not technically a pilot program or demonstration project, was

intended to serve as a blueprint for future similar Con Edison demand-focused DSM programs. As such, the final phase of the program may have been the most important for informing future program design. A phased data collection effort would have addressed ongoing changes to the program, especially the program in its final form. While the exact nature of the program changes and external factors may not have been predictable, a phased data collection effort would have allowed for a more dynamic view of the program over time and captured the program in its final, most mature form.

With a phased data collection approach, research topics can also be adjusted to assess program changes over time. For example, participants can characterize changes to their satisfaction with the program during different "phases" or iterations. Such questions help support the need for summative reporting (often required by state regulatory entities) within a continuous improvement environment by characterizing the customer perspective on whether the program improved over time. On the other hand, this approach can also support prospective evaluation by comparing the final phase of the program to earlier phases. That is, it helps answer the question of whether the program in its final form was more or less effective than in earlier phases.

Reporting

Within the CIP context, interim reporting is critical to success as it provides the means to communicate evaluation findings as quickly as possible. What may be less obvious is that summative reporting can also serve as an important tool. While not facilitating continuous improvements, summative reports can be used as a tool to comprehensively present evaluation results, which utilities can use for regulatory purposes.

In the case of the DMP, presenting interim results to implementers was critical for program staff to better understand evaluation results as early as possible. The evaluation team presented interim results overtime through a staging of data collection efforts. Because evaluation activities were staged in an intentional way, the evaluation team could report on specific activities throughout the course of the evaluation. Because these activities were designed to cover initial stages of the evaluation, the evaluation team could report with confidence on what the results meant. The evaluation team shared these results through written memos with the utility.

In addition to interim findings, the evaluation team provided a summative report. Many jurisdictions require utilities, like Con Edison, to provide summative evaluation reports. These reports provide regulators and other stakeholders insights into whether program funds were spent efficiently and according to plans. They also provide recommendations for future program plans based on evaluation findings. This was true for Con Edison whereby evaluators needed to provide the utility and regulators with a report that presented program activities and evaluation findings. Developing a summative report, however, posed challenges to the evaluation team since it needed to describe the dynamic nature of the program, while simultaneously presenting actionable recommendations above and beyond recommendations that program staff had already addressed. The evaluators implemented two strategies to address this challenge.

First, the evaluation team defined program activities, goals, and objectives. Since original program documentation was outdated, the program description relied on primary data that it collected from program staff throughout the course of the evaluation. This meant that these conversations were well documented to provide defensible findings within a regulatory context. It also meant that the evaluation team vetted program changes that occurred throughout the year to make sure it was accurately representing the program in the final report. Lastly, to ensure that the final report truly reflected the program, evaluators met with program staff to review the draft report and worked closely together to make sure the report accurately presented any aspect of the program that might have changed throughout the course of implementation and reasons why those changes occurred.

Second, because the implementation team incorporated evaluation recommendations into their program design throughout the course of the evaluation period, the utility addressed a substantial portion of the evaluation recommendations prior to presenting the evaluation report to their regulatory body. To account for this fact, the evaluation team included descriptions of how the utility implemented program recommendations, ensuring that concerns or red flags raised by the identification of the issue were adequately contextualized for regulators and future audiences. The summative report also included interim memorandums from the rapid feedback research, thereby formally documenting the findings and recommendations that the evaluation team communicated throughout the course of the evaluation.

Conclusion

CIPs, like the DMP, are becoming more common and programs such as these require more rapid feedback on potential implementation issues simply because failure to hit program goals carries with it severe outcomes (i.e. potential for system failure). This increased need for rapid feedback evaluation findings to support program implementation requires evaluation agility and closer contact with the program implementation team than past summative evaluation approaches necessitated. By applying the following three lessons, evaluators can be more confident that they can provide valuable feedback to program implementers, while implementers are able to quickly improve program processes and are therefore better able to position themselves to meet and/or exceed their goals.

- **Communication –** Evaluators and program implementers need to develop a collaborative relationship with frequent and well documented conversations. A collaborative relationship facilitates information sharing from program implementers to inform evaluators of program changes, and preliminary results sharing from evaluators to inform program implementers with early evaluation result themes or recommendations. Such open and frequent dialogue can be new to many program staff and therefore the authors recommend engaging with staff early in the evaluation process to best understand communication needs and protocols. Protocols could include creating a document that can track program changes, thereby creating a reference tool to frame these discussions.
- **Data Collection –** Effective data collection in a dynamic environment is grounded in a collaborative communication environment. In this space, evaluators can phase data collection based on current and planned program activities. Sampling approaches, such as phased waves, may better account for known and unknown program changes that might occur, and assure that the evaluation captures feedback on the program in its final, most mature iteration.
- **Reporting –** While there is still a need, oftentimes, for a summative evaluation report, such a report needs to tell the story of program changes that occurred and explain why these changes occurred. The report can also present evaluation recommendations that were made through the evaluation period but highlight how program staff addressed these recommendations.