SEARCHING FOR A FINE LINE…
IMPLEMENTERS AND EVALUATORS "TOGETHER, BUT NOT AS PARTNERS"

Jennifer Mitchell-Jackson, Opinion Dynamics Corporation
Brad Kates, Opinion Dynamics Corporation

Introduction

This paper explores the balance between evaluation research and energy efficiency program implementation. Through this paper, we meld our experience and perspectives with the comments of several program implementers from across the country. Our hope is to spark new ideas, and to start a discussion that will ultimately increase the value of evaluation research.

While conducting research for this paper, we spoke with thirteen program implementers. While not comprehensive, we sought out comments from three types of program implementers:
  ➢ Utility program implementers
  ➢ State-run or regional organization program implementers
  ➢ Contractors.
The thirteen implementers included five utility program implementers, three program implementers from regional or statewide organizations, and five implementing contractors. In addition to representing three different types of organizations, these thirteen also represented eight different states including Texas, Wisconsin, Oregon, California, Vermont, Massachusetts, New York, and New Jersey.

Although tempered by the fact that we have been working in the evaluation field, this paper represents a one-sided perspective since it presents only the findings from discussions with program implementers. Future research will include similar interviews with utility evaluators and evaluation contractors. While not comprehensive nor quantitative, we seek to use our interviews with program implementers as a springboard for a larger discussion.

It should also be noted that the anecdotes that we weave through this paper are not meant to generalize all evaluations. They are only shared to raise awareness of particular models so that even better models can be created and realized.

While this paper does not tease apart the differences between the various types of evaluation research (such as impact evaluations, process evaluations, baseline studies and market assessments), it does provide a general perspective on the field of evaluation research—as seen by program implementers.

Background

On the surface, the implementers that we spoke with generally find evaluation research to be useful, and most are pleased with the process involved in conducting evaluation research. As we inquired more deeply, however, many were only vaguely familiar with the evaluations that have been conducted on their programs. Except in a few cases, implementers are brought into evaluation research fairly minimally. They are asked to provide data, and they may be interviewed, and then most (although not all) receive a copy of the final evaluation report.

Comments from several of the implementers that we spoke with seemed to indicate that evaluations are not particularly useful in their current form. In the opinion of one implementer, "evaluation is just to confirm—in a statistically valid and unbiased fashion—what implementers already know."

From discussions with implementers, there were several areas of concern that were recurring. These included:
- ➢ Obsolete evaluation protocols,
- ➢ Isolated recommendations,
- ➢ Unreasonable data requests,
- ➢ Less than valuable reporting, and
- ➢ Lagging evaluation schedules.

These concerns are described in more detail below.

Obsolete Evaluation Protocols

Evaluation is lagging behind program implementation, according to comments from implementers. Evaluations are most often based on the model of a utility running a demand-side management program. The assumption is that energy efficiency programs need a traffic cop.

Today, however, the goals of many utility programs have changed (e.g., many are working towards transforming the market rather than just reaching a certain kW savings). Furthermore, there are several non-utility organizations—such as regional or statewide agencies—that are implementing energy efficiency programs. Thus, very different entities with very different goals are often being evaluated under the same model.

The prevailing comment of representatives from state and regional organizations is that it is their job to make sure that the programs are meeting the needs of their communities. According to one respondent, he sees himself as a "public employee, spending public funds, trying to capture public benefits for public good. [He has] a completely different motivation." These implementers are interested in evaluation research that can assist in program design and inform program development, rather than the policing of their program.

On a similar note, one implementer pointed out that evaluation is often stuck using definitions and protocols that are no longer practical. For example, in a recent C&I evaluation, the definition of a "freerider" included any customer that had previously been given an incentive payment. According to the protocol, this definition was supposed to eliminate freeridership. Program goals, however, often state that programs should target owners of multiple buildings in an effort to have a greater effect.

Despite these comments, there are a few implementers who acknowledge that not all evaluation efforts are completely static. Some evaluators are moving towards more robust evaluations to assist in program design and inform program development. Furthermore, one implementer that we spoke with was impressed with some of the market-based information that he received from the evaluation on his program.

As program implementation models and goals advance, evaluation needs to keep in step. It may be necessary to distinguish between "regulatory evaluations" and more "actionable" evaluations, which are meant to inform program design. Furthermore, evaluation protocols need to be flexible enough so that they can adapt to the needs of the program.

Isolated Recommendations

To be valuable, evaluations need to be put into the context of the program. Most often, the evaluation contractor enters the scene at the end of a program year and has no firsthand knowledge of the program. While most of the selected evaluators have read a program description and have some general knowledge about the markets that are served, they often do not know about the intricacies of the program. They generally learn about programs from literature that is dated and cryptic. While evaluation methodologies often allow for depth interviews with implementation staff, these 30 min to one-hour interviews do not provide enough time to recount all of the details of the program. Implementers point out that, without context, evaluation findings are often misconstrued.

Improved communication is already occurring in some evaluations. Several implementers mentioned that they are often provided with a rough draft of the evaluation findings and are asked to comment prior to finalization of the report. Through this process, implementers are able to provide context to some of the findings and to correct any details that may be inaccurate. Implementers also recognize that it is valuable to get a contractor that really knows the market and/or really knows the program. For this reason, in an ongoing program, the incumbent evaluator is often selected.

In addition to ensuring that the evaluator is in-tune with the needs of the program, the utility implementer or implementing contractor should have the opportunity to inform the evaluation process. One implementer recalled an evaluation effort where they were not brought into the process. The evaluator conducted a survey to assess consumers' association of a local energy efficiency offering with the local utility. The program, however, was not charged with branding the service to the utility's name, so the information collected through the evaluation was not valuable.

Although implementers provide the evaluator with data, the implementer usually does not know how the evaluator will use this data. For example, one implementer provided detailed data about the retailers involved in their energy efficiency program. The evaluator then used this information to conduct mystery-shopping visits to the retailers in order to assess the number of energy efficient products carried. Unfortunately, however, the visits were performed at a time when many stores had not yet received their shipments of the products due to delays caused by the program.

The research community has been working to tie evaluation efforts into program logic. The programs themselves, however, have not always clearly laid out or updated their program goals. Moreover, some evaluations are misguided at the direction of the Request For Proposals (RFP). One recent evaluation was conducted to meet the objectives laid out in the RFP. The evaluation dealt with the coupon component of the program, but the program had been recently changed to reach out beyond the coupon process. The program also contained a retailer initiative; however, this component of the program was only uncovered in the depth interviewing, which occurred after the final evaluation workplan for the evaluation was complete.

It may be that while the programs are judged on certain metrics (for regulatory purposes), evaluating based on these metrics does not always assess the overall goals of the program. Implementers need to clearly lay out and periodically update the goals of their program, and these goals need to be shared with the evaluation contractor.

Ideally, evaluators should be involved upfront. In the case where there is a *utility* or *energy efficiency organization* evaluator, this person should be tracking the program so that they are knowledgeable about it at the outset of the evaluation. Alternatively, the evaluation *contractor* could be brought on at the

beginning of the program year to periodically check in. Energy efficiency programs should also consider hiring evaluation contractors for the program rather than for individual projects. This will allow the evaluation contractor to become more familiar with the program.

In addition, it may be valuable to meet with implementers at the time of the evaluation, or at least inform implementers of the scope of work so that they are tied into the process more. Ideally, the scope of work for the evaluation would be determined prior to the program year so that the implementers understand how it will be evaluated.

Unreasonable Data Requests

Implementers frequently mentioned data requests from evaluators as an area of concern. According to implementers, evaluators make requests for data, which potentially could have serious impacts on the program. Some data requests can preclude customers from participating, while other requests are impossible to meet. The most frequent comments were that "evaluator's data requests are often unrealistic and time consuming."

Implementers mentioned that it often takes a couple of days to meet the data request since they usually have to query and correct data. They also mentioned that there is not a lot of room in their budgets to meet these requests. Although they sometimes renegotiate, usually the implementer just subsumes this cost into their program budget.

Furthermore, data requests that are made after the program has been implemented—particularly if new data requests occur repeatedly—can undermine the program. Participating retailers and manufacturers may feel that it is no longer worth the trouble, so participation may dwindle. In this and many other cases, data requests begin to pose a barrier to participation in the programs.

In addition, program implementers do not always keep files and schedules of all marketing efforts and materials. When the evaluation begins, therefore, implementers scramble to collect this information, and important pieces are sometimes overlooked.

Ideally, since evaluation is a necessary part of the program, the program should be designed to collect marketing and educational materials along the way. Data collection requirements should be determined during the program development phase (or at the beginning of the program year). The best models of this bring the evaluator in at the beginning of the program year so that they can provide input on what information should be collected. Acting early on also allows implementers to negotiate about the type of information necessary so that data collection is not a barrier to participation.

Implementers noted, however, that it is best to have the utility or energy efficiency evaluation staff acting as the intermediary between the implementing contractor and the evaluation contractor. This model will help to keep both parties within the scope of their contracts and help to eliminate unnecessary data requests. It will also help to maintain the integrity of the study.

Less Than Valuable Reporting

In general, 3-10% of the program budget is spent on research; yet, oftentimes the research gathered through the evaluation process never reaches the program implementers. Alternatively, it may be handed to the implementer but then stuck on the shelf.

For the most part, comments on the reporting indicated that there was too little information provided to implementers, and most of it is in a format that is not useful to the program.

According to implementers, there are things that evaluators can do to increase the value of the evaluation research, such as pulling out relevant material and creating one- or two-page summaries that can be used by the implementers to move the market. Market-based information can be extremely valuable to the program. Depending on the program, information on market penetration and/or freeridership and spillover may be able to be repackaged and used to encourage manufacturer participation in the program.

Furthermore, although implementers seemed to indicated that they receive the draft evaluation report more and more often, there are times when they receive only the final report or do not receive the report at all. And even respondents who were given drafts of the evaluation seem to indicate that this process is inconsistent. There is no clear process for getting implementers' feedback.

Because of this, implementers acknowledged that the recommendations are often drawn in a vacuum and are not always practical. While they may not be held to acting upon the recommendations, the value of the evaluation is reduced. In addition, interveners may come in at a later date and see the report and inquire why the evaluator's recommendations were not acted upon.

Implementers, therefore, would prefer quicker feedback loops and results that are quantified to be more useful to managers, such as summary reports highlighting elements for program participants. It would also be useful for evaluators to interview participants right after they have participated, rather than several months later. This model mirrors some of the utility transactional surveying and could provide immediate and actionable feedback. Furthermore, it would ensure that participant feedback and freeridership information is more accurate.

Lagging Evaluation Schedules

Several implementers felt that evaluations were "not timely." Evaluation findings are often finalized after the next program year has started, thus the timing of the evaluation research is usually out of sync with the program cycle. One implementer commented that he thought that this was by design. He explained that evaluation research takes a longer-term perspective; it is not "a real-time on the ground how are we doing?" Rather, the timing of evaluation research is delayed and retrospective: "11 months after the fact, I get a report that customers like working with the program, but I already know that they like working with the program." Because the evaluation research is not synchronized with the program, implementers can make minor tweaks, but often not bigger changes.

To increase the value of evaluation research so that the program implementers can incorporate the results into their program planning, it is incumbent on the program to get RFPs out early enough to meet these deadlines. Implementers should be informed of the timing of the evaluation and be expecting data.

Summary of Findings

While there is no one "right" model of conducting evaluation research, any model needs to include frequent communication between implementers and evaluators. There needs to be separation between evaluator and implementers but to maximize the value of evaluation research, it should be a fine line. A successful evaluation needs to have someone who understands the details of the program before the evaluation is started. Furthermore, the evaluation may be made more useful if there is an implementer who is able to provide insight along the way.

Specific recommendations include:
- ➢ Making evaluations adaptable to the program needs
- ➢ Setting program goals upfront
- ➢ Getting RFPs out early so that the evaluator can be brought in early and the evaluation will be on the program cycle
- ➢ Hiring evaluation contractors for the program rather than for specific projects
- ➢ Working together to determine data collection pre-program
- ➢ Exposing implementers to recommendations as they are formulated
- ➢ Asking implementers to review evaluations
- ➢ Presenting information in a useful format to the program.

There are some models, such as having the evaluation contractor working directly for the implementing contractor, which do not work. In these cases, the integrity of the study is compromised because the implementer is overseeing the evaluation. However, there is a need for increased communication.

Acknowledging the increased communication between the evaluator and the implementer can lead to more valuable evaluation research. Once that is done, the next step is realizing that this will incur both time and money. Implementation budgets, therefore, should consider building in a cushion for evaluation-related tasks such as:
- ➢ A sit-down meeting with the evaluator at the beginning of the project
- ➢ Responding to data requests
- ➢ A final debrief meeting with the evaluation contractor prior to finalization of the report.

The best process is to determine upfront what program success entails. If the goals are clearly laid out, evaluators can then measure the program against these goals. In addition, the evaluator can be brought in at the beginning of the program year to help determine data collection requirements and to lay out a scope of work for the evaluation. This reinforces the goals of the program and helps to keep the implementer on track. In this way, implementers can help shape evaluation without having too much control.

Of course, there are also drawbacks to bringing implementers into evaluation. Adding an additional layer of review may hold up the evaluation process. This can be dealt with by providing the implementer with a deadline for comment. If no comments are received, at least the information has reached the implementers. Additionally, as implementers become more involved, some caution should be taken to ensure that program implementers do not dictate where evaluation funds are spent.

Ultimately, however, the design of the evaluation effort needs to be incorporated upfront into the design of the program.