

Perfect Strangers: Evaluators Working with Implementers

Christopher Dyson, DNV GL
Chere Coleman, Michigan Community Action
Tim Hennessy, DNV GL
Ben Jones, DNV GL

ABSTRACT

One of the biggest challenges in energy efficiency program evaluation is providing evaluation findings which can influence program design and delivery in a timely manner. Program implementers often state that evaluation findings come too late. In some cases, the evaluation findings point to issues the implementers have already addressed. In other cases, the findings are still relevant but are difficult to implement because program designs and delivery mechanisms are already set in place.

This paper discusses how program evaluators and implementers can overcome these problems through closer collaboration and faster delivery of evaluation findings. We describe practices from our evaluation work in Michigan including: 1) providing early feedback on proposed energy efficiency projects to help screen out projects that risk lower realization rates; 2) performing quarterly reviews of program tracking databases to provide "early warnings" on issues that could put the program at risk of not meeting savings goals; and 3) regular meetings between evaluators and implementers to discuss issues of common concern.

The paper will also discuss issues that can make evaluator/implementer collaboration more challenging. These include inadequate levels of trust between evaluators and implementers and the need for program evaluators to preserve their independence and objectivity.

Key Issues in Program Evaluator and Implementer Interactions

This section will frame some of the key issues to consider when examining program evaluator and implementer interactions. It will also touch briefly upon another focus of this paper: the importance of timely evaluation results.

The Common Interest of Program Evaluators and Implementers

Historically, the relationships between energy efficiency program evaluators and implementers have rarely been collaborative ones. The co-authors of this paper, who have over 50 years of cumulative program evaluation experience, have seldom witnessed such collaborations. In contrast, the relationships between program evaluators and implementers are usually characterized by mutual distrust and disappointment. Occasionally these relationships will even deteriorate into outright antagonism.

There are many reasons for these difficult relationships. However, it is necessary to first take a step back and recognize that program evaluators and implementers do have an important common interest: the success of the energy efficiency program.

An unsuccessful program can have many negative consequences including unhappy program participants, missed opportunities for energy savings, high levels of free ridership, the

loss of performance incentives for utilities or implementation contractors, increased regulatory scrutiny, the replacement of the implementation contractor and program termination.

While these negative consequences often weigh more heavily on program implementers, they can also create problems for program evaluators. For example, if an impact evaluation finds that a program saved much less energy than was originally estimated; evaluators will likely face resistance and increased scrutiny of their findings and methods. The strongest resistance will usually come from those who benefit most directly from the program achieving its savings goals, such as implementation contractors and utilities. Yet resistance to such evaluation results may also come from regulators and system planners who may have made key regulatory decisions or planning assumptions based on pre-evaluation estimates. Whenever there is this much resistance to negative evaluation results, there is always a risk that it might bias the evaluation in the direction of telling a more positive “story” about the program than might be otherwise warranted.

Accordingly, a successful energy efficiency program benefits both program evaluators and implementers. Therefore it is in their mutual best interest to work together to increase the chance of program success. The next section discusses why this so rarely happens.

The Seeds of Discontent

The relationships between program evaluators and implementers are often fraught with distrust and disappointment and can sometimes become antagonistic. One common reason is that program evaluators are often the bearers of bad news. For example, post-evaluation program energy savings estimates are often lower than pre-evaluation program energy savings estimates. Therefore it is not surprising that many implementers welcome impact evaluators with the same enthusiasm as businessmen receiving visits from IRS auditors.

In theory, program implementers should view process evaluators more favorably, since process evaluations often suggest ways to increase program participation and improve operational efficiency. Yet program implementers often say that process evaluation findings are provided too late to be beneficial. In some cases, the evaluation findings point to issues that the implementers have already addressed. In other cases, the findings are still relevant but are difficult to implement because program designs and delivery mechanisms are already set in place. Finally, some program implementers may view the process evaluation’s identification of program shortcomings as the equivalent of saying that “their baby is ugly.”

Whenever one party is assessing the performance of another, there will be some tension between the assessor and the person being assessed, whether the evaluators are delivering good news or bad news. This is true in all aspects of life: job performance assessments, restaurant reviews, gymnastics competitions, etc, as well as efficiency program evaluations.

Lack of frequent interaction between program evaluators and implementers can also exacerbate any inherent tensions in these relationships. Many program evaluations do have frequent interactions between program evaluators and implementers, but often within the limited context of data exchange. For example, the person on the evaluation team who designs the sample for participant surveys will often communicate frequently with the person on the implementation team who manages the program tracking data. It is possible that this working together can cause them to develop a good rapport. Yet, in our experience, it is rare that there will be frequent interactions between program evaluators and implementers at the higher management levels. More commonly, the program administrator or contract manager – often a utility – will have separate meetings with the program evaluation and program implementation

teams. Figure 1 shows the typical management and communication structure for energy efficiency programs where the direct communications between program evaluators and implementers are very limited and only occurring in the context of data exchanges.

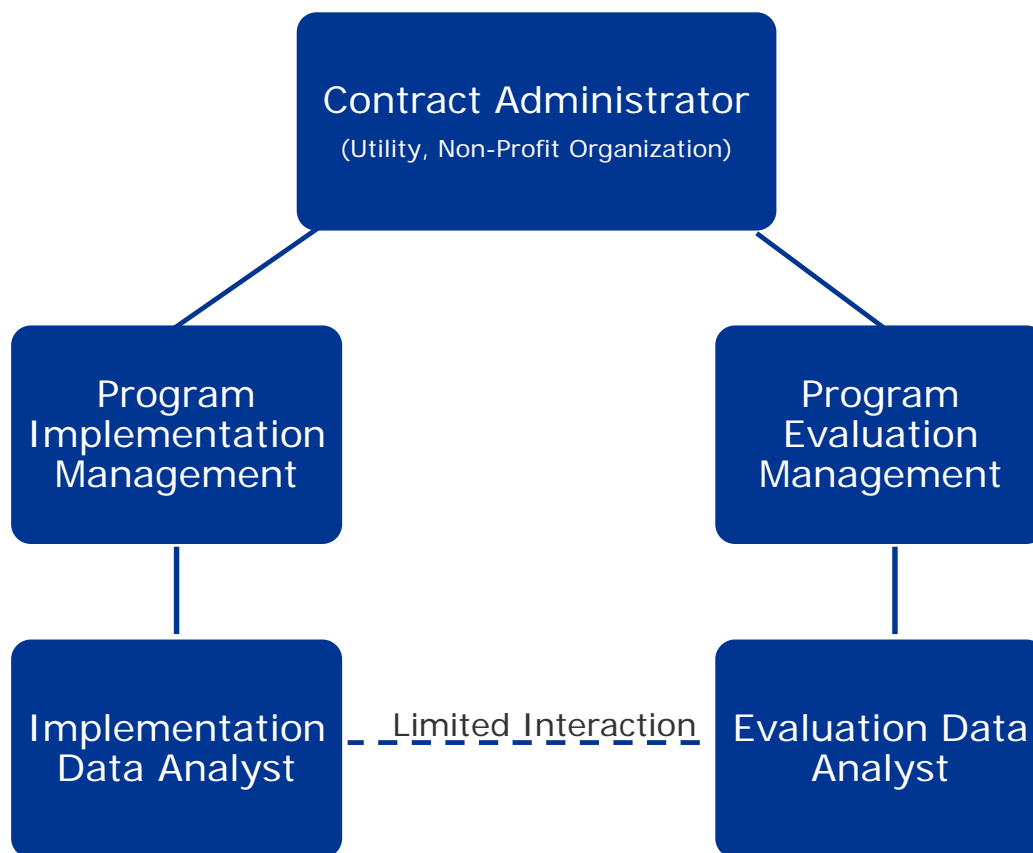


Figure 1. Typical Management and Communication Structure for Energy Efficiency Programs

The next subsection discusses some of the risks of lack of interaction between program evaluators and implementers and the benefits of greater interaction.

Knowledge, Empathy and Trust

Lack of interaction between program evaluators and implementers, especially at the higher levels, can lead to lack of knowledge, empathy and trust between the two groups. Lack of knowledge -- not knowing what the other side is doing and why they are doing it -- can create difficulties on both ends. For example, a process evaluation may find that one of the program rebate offerings is too low. If the program evaluators are unaware of the program's deadline for updating the rebate application forms, they may issue this finding after the opportunity to implement the rebate increase has passed. Similarly, if program implementers are unaware that the program evaluators are going to interview installation contractors, they may not be very thorough in collecting contact information from these contractors.

Empathy involves identifying with and understanding another's situation, feelings, and motives. When program evaluators and implementers interact more frequently, they get to know

their counterparts not just as names in an email; but also as voices, faces, and personalities. This increased familiarity makes it more likely that they will be attuned to the situations, feelings, and motives of their counterparts. So their knowledge will extend beyond simple facts – e.g., that a program evaluator or implementer has a deadline – to the deeper awareness of how important this deadline is to their counterpart and the level of anxiety and stress that would result if the deadline was missed. This empathetic awareness should give them a greater sense of responsibility, and even urgency, to help out their counterpart (e.g., provide important data in a timely manner) than if their knowledge of this deadline was purely factual.

Developing trust between program evaluators and implementers is difficult. It not only requires closer interactions between the two parties, but also a track record of openness and fair play between them. Things inevitably go wrong in both program delivery and evaluation. If the evaluators and implementers do not trust each other, they will likely conceal these problems from each other. Implementers may fear that revealing a problem with their program delivery will lead to negative findings in a process evaluation report. Evaluators may fear that revealing a problem with their evaluation will give implementers ammunition to discredit the findings.

Yet often the concealment of these problems benefits neither party. Smaller problems with program delivery, if uncorrected, may lead to bigger problems in the future. For example, if a C&I program is not screening potential energy efficiency projects for free ridership, the end result could be low net savings for the program. In addition, the concealment of these problems can mean a missed opportunity for a solution from the other party. For example, program evaluators have experience recognizing characteristics of energy efficiency projects which give them a high risk for free ridership. So if a program implementer has had problems with high levels of free ridership in the past, it would make sense for them to engage the program evaluators to help them find better approaches for filtering out projects with a high free ridership risk. However, in our extensive program evaluation experience this rarely happens. The main reason it rarely happens is because the program implementers and evaluators do not trust each other enough to be willing to reveal their problems to the other party and seek solutions.

Of course, there are no guarantees that closer interactions between program evaluators and implementers will lead to better relationships and better program results. It is possible to envision scenarios where serious disagreements between evaluators and implementers as to the best ways to estimate energy savings or free ridership could be made worse by frequent interactions. However, as this paper relates, our recent experiences in Michigan indicate that more frequent interactions between program evaluators and implementers are, on balance, an important contributor to program success.

The Role of the Contract Administrator in Evaluator/Implementer Interaction

The administrator of the program evaluation and implementation contracts, which is usually a utility or non-profit organization, plays an important role in influencing the interactions between program evaluators and implementers. These contract administrators usually decide whether the program evaluation and implementation teams meet together or separately. They also serve as a filtering mechanism for any remote (email, phone) communications between evaluators and implementers.¹ For example, when evaluators provide information which may be

¹ These communication protocols often stem from the structure of the contractual relationships. Usually the energy efficiency contract is between the state governmental or regulatory entity and the energy efficiency contract administrator and the program evaluators and implementer have no contractual relationship.

of interest to implementers, it may get to them in unfiltered form (e.g., a forwarded email). Yet the contract administrators may choose to convey the information to the implementers in altered form or may choose not to share the information at all. Finally, the contract administrators often serve as arbiters for any disagreements between evaluators and the implementers.

There are a number of possible reasons why contract administrators may want to keep program evaluators and implementers apart. Some may believe that keeping them apart will enhance the appearance of the evaluators as being independent, third-party assessors of program performance by reducing the risk of cooption and keeping the relationship at “arm’s length.” This may be especially important in cases where the contract administrators may be under greater scrutiny from regulators and other stakeholders about possible conflicts of interest. For example, if the contract administrator is a utility which is also eligible for financial incentives for good program performance, that utility might keep the program evaluators and implementers separate to enhance this appearance of evaluator independence.

Contract administrators may also want to keep evaluators and implementers apart because they view their own roles as necessary “translators” between two parties who speak different languages. As noted, program evaluators and implementers may have very different perspectives as to how energy savings or free ridership should be calculated. They may also disagree as to the best focus of any market research. Some contract administrators may believe that bringing together parties who have such different perspectives will only lead to confusion and possible antagonism.

However, there are also good reasons why contract administrators may want to bring program evaluators and implementers into closer interaction. We will discuss these reasons in our case study.

The Importance of Timely Evaluation Results

In the delivery cycle of every energy efficiency program, there are periods of time when it is possible to change elements of program delivery and design without too much disruption. Yet these windows of opportunity are very brief. At a certain point, the program needs to finalize its rebate application forms and marketing materials and communicate the program requirements to its trade ally network.

Because these opportunities for altering program design and delivery are so fleeting, it is very important that the program evaluation team can produce findings and recommendations in a timely manner. For these findings and recommendations to be “actionable,” they must be available when it is feasible for the program to take action.

Yet most program evaluations follow a cycle in which the program evaluation results are available after the program year has ended. The top half of Figure 2 shows a traditional evaluation schedule where program year 1 evaluation results only become available after the implementation of the year 2 program cycle. For programs which do not change much from year to year, evaluation findings and recommendations based on the previous program cycle can still

Accordingly, many administrators insist on being notified of all contact between the program evaluators and implementers.

be useful for informing future program design and delivery.² For programs which are rapidly changing, however, evaluation findings and recommendations based on older program designs have limited usefulness.

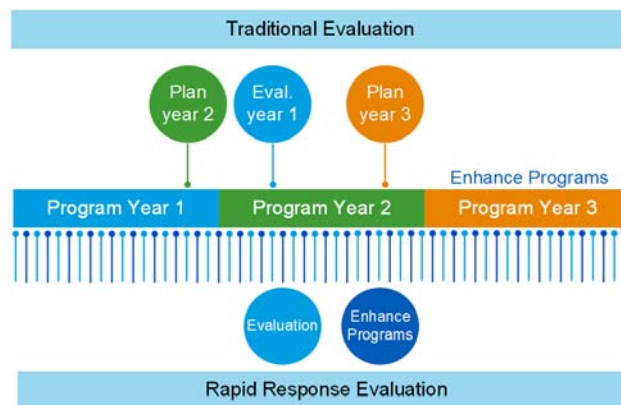


Figure 2. Traditional Evaluation Scheduling vs. Rapid Response Evaluation

A Case Study: Michigan Energy Optimization Programs

Our experience evaluating various Michigan energy efficiency program portfolios provides a useful case study for interaction and collaboration between program evaluators and implementers. This section begins with some necessary background information about the key stakeholders and the regulatory context. It then describes this interaction and collaboration between evaluators and implementers and the lessons we learned from it.

The Regulatory Context and the Key Stakeholders

Michigan’s 2008 Public Act 295 requires all utilities in the State to design and implement a comprehensive set of Energy Optimization (EO) programs. Utilities who are unable or unwilling to implement EO programs on their own, or who wish to take advantage of economies of scale cost savings, can meet the EO requirements by participating in a utility collaborative which administers a portfolio of energy efficiency programs. These Michigan collaboratives include Efficiency United (EU), the Michigan Public Power Association (MPPA), and the Michigan Electric Cooperative Association (MECA). Section 91 of the legislation allows the Michigan Public Service Commission (MPSC) to select qualified independent non-profit organization to manage the contracts for these utilities.³

² It is important to note that this paper is not advocating that timely evaluation findings replace annual reviews but rather that they complement them. State regulations may require that program’s energy savings accomplishment be measured on an annual basis to determine the program administrator’s or implementer’s eligibility for performance incentives. There is also some value in producing a complete assessment of a program’s yearly performance, just as there is value in producing a complete assessment of a company’s yearly performance in the form of an annual report.

³ Although these collaboratives share some similarities, they are not alike in terms of their regulatory status. EU is a utility collaborative for which the MPSC selected a non-profit entity (MCA) to be the contract administrator for the implementation of energy efficiency programs in the service territories of these

Act 295 requires that the utilities achieve a target gross energy savings that had been independently verified by an Evaluation Contractor (EU). The statute also asks for annual regulatory filings with the MPSC which document the claimed and verified savings. The EC is responsible for independently verifying the claimed savings by the Implementation Contractor (IC) on behalf of the client, certify the savings and prepare a document for the client for filing with the MPSC. DNV GL is the independent Evaluation Contractor (EC) for 51 utilities throughout Michigan under multiple contracts. The main focus for this case study will be the evaluation of the energy efficiency programs under the contract for the EU collaborative. Currently 18 utilities participate in the EU collaborative.

The Michigan Community Action Agency Association (MCA) is the non-profit which manages the program evaluation and implementation contracts for the EU portfolio of energy efficiency programs. CLEAResult Consulting is the main implementation contractor for this EU portfolio. It does so under its own contract with the MPSC.

How the Collaboration between Program Evaluators and Implementers Began

DNV GL began its evaluation of the Michigan energy efficiency programs in 2009 soon after their first year of program implementation. In the first few years of our evaluation, the relationship between DNV GL and the implementation contractors was traditional. The direct in-person interactions were typically limited to the annual kickoff meetings and “wrap-up” meetings. At the kickoff meetings, the implementation team would describe the program designs and the evaluation team would describe the initial evaluation plan. In the wrap-up meetings, the evaluation team would present the final verification results. The only direct communications between the evaluators and the implementers involved data queries concerning the program tracking databases. All other communications were filtered through the contract administrator.

However, over time, the relationship and the interaction between program evaluators and implementers began to evolve. First the MPSC started to convene meetings of interested parties to facilitate the energy efficiency objectives of the EO statute by giving a forum to develop common understanding of issues and address common challenges. These meetings provided the initial step in changing the traditional way that the evaluators and implementers interacted. They offered a less formal environment that allowed the frank discussion of complications related to the implementation and evaluation of a full spectrum of programs offered by the utilities.

As the implementation of energy efficiency program evolved, practical considerations necessitated more interaction between the implementers, evaluators, contract administrators and regulators. Traditionally the program evaluators had not published the verified savings until after the conclusion of the program year. Since the ultimate verified and certified savings were the basis of the compliance determination, this schedule introduced a certain level of uncertainty for the program implementers and contract administrators.

Early Desk Reviews

To reduce the uncertainty associated with ex post de-rating of claimed energy savings, the contract administrators began to informally request the evaluators’ opinions on the viability of certain proposed energy efficiency project opportunities during the implementation year.

utilities. The other utility collaboratives mentioned (MPPA and MECA) formed independently, out of mutual self-interest, and had no direction or sanction from the MPSC.

These requests included querying the evaluators on the appropriateness of proposed engineering approaches to determine savings, or the assignment of savings to certain customer classes.

One example of a request for an early evaluator project assessment concerned a very large energy efficiency project for a Michigan manufacturing plant. The project was complex, and ultimately would represent a large portion of the contract administrator's annual claimed savings. If the project was disallowed or significantly de-rated during the verification process, the contract administrator would not have met the statutory energy savings requirements.

Typically the gross savings estimation of such a large and complex project would take place in two stages. In the first stage, an experienced engineer on the evaluation team would examine the project documentation and energy savings calculations to verify that the savings estimates were reasonable and derived using commonly-accepted methods. In the second stage, an evaluation engineer would visit the site to verify that the measures were installed and operating as intended. The evaluators would then use the information collected from these two stages to determine the amount of verified claimed savings.

In the case of this manufacturing plant project, the contract administrator asked DNV GL, as the evaluation contractor, to conduct the first stage of the gross savings verification on this project (the engineering "desk review") before they provided any program incentives to the plant. If the evaluators were able to verify that the savings estimates were reasonable and based on commonly-accepted methods, this would significantly reduce the risk of the onsite verification later de-rating the energy savings. Since this request only changed the timing of the engineering review, it did not compromise the independence of the program evaluators.

This early review process was successful. The contract administrator and program implementer were able to move forward with the large project after receiving the findings from the early engineering desk review. Based on this positive experience, the contract administrator requested that this process be a formal part of the evaluation process. It was understood that this early warning procedure would only be used for: 1) very large projects where the de-rating of the energy savings would significantly impact program savings achievements; or 2) projects which had great complexity or unusual elements that would increase the uncertainty of the pre-evaluation energy savings estimates.

How Communications Evolved

Communication is the foundation upon which a healthy evaluator-implementer relationship is based. Without open lines of communication, trust breaks down and the risk of expensive misunderstandings increases. With better communications, implementation teams are able to see the value which independent evaluation provides. Better communications also allow evaluation teams to know earlier about program implementation practices which might increase evaluation risk down the road.

Trust takes time. For the first three years (2009-2012) of the evaluation of the EU program portfolio, direct interaction between the program evaluator (DNV GL) and the program implementer (CLEAResult) was very limited. During this period, interactions were limited to annual kickoff and wrap-up meetings, data and document requests, evaluation report reviews, annual program manager interviews and a handful of ad hoc interactions to resolve data issues. This infrequent and relatively distant interaction did nothing to dispel the natural wariness that can develop between evaluators and implementers, as discussed earlier in this paper.

One thing that helped change the nature of this relationship was EU's introduction of a group of Special Pilot programs in 2013. Because many of these program designs were relatively untested, MCA, the EU contract administrator, requested that the evaluation and implementation teams attend monthly meetings together to learn about and document the programs as they developed and changed. As part of the documentation, the evaluation team provided near-real-time feedback and recommendations on how to address program challenges. Typically these recommendations would start as an idea expressed at the monthly meeting which the evaluators would then analyze and expand upon in the monthly memo.

In these monthly meetings, the evaluation and implementation teams learned to "lower their guard" and discuss challenges and ideas openly and frankly in order to find solutions. This discussion often took the form of evaluators asking the implementers questions about interim program achievements, current program strategies and their track records, and future program design adjustments and implementation strategies.

The exchange of ideas allowed both sides to see each other as being engaged and interested in improving program performance. There was also mutual benefit in a clearer understanding of program objectives. By hearing from the implementers what the program objectives were and the thinking behind them, the evaluators could better assess whether the program had met its objectives. This also benefitted the implementers because they had greater assurance that evaluators were assessing the program based on its true objectives, rather than less-informed assumptions about the program objectives which might not be accurate.

These monthly discussions happened in the presence of the contract manager (MCA) and the regulator (MPSC). One advantage of MCA as a contract manager is that they were a non-profit organization which did not benefit as directly from the program energy savings achievements as would be the case for a utility with performance-based incentives. This allowed MCA to better serve the role of "neutral mediator." As noted, one concern about closer interaction between program evaluators and implementers is that there might be a chance of cooption or collusion between the two parties. Having both the presence of contract manager and the regulator in the room helped to improve the transparency of the process.

The benefits of having the MCA and MPSC in these monthly meetings went beyond their roles as neutral mediators and regulatory "watchdogs." First, they were a fertile source for ideas on ways to improve energy efficiency design and delivery. Second, MCA's and MPSC's involvement in the discussion gave them some ownership over decisions to change program implementation or evaluation strategies based on interim program results. This led to quicker approvals of any such changes.

For example, one monthly meeting discussion focused on why so few commercial and industrial (C&I) customers were taking advantage of the EU commercial financing program. DNV GL suggested that a C&I general population phone survey, which was not part of the original evaluation scope, would help reveal what financing mechanisms these C&I customers used and preferred and whether they were aware of the program financing options. MCA and MPSC both quickly approved this survey because both had been involved in the discussions which revealed its usefulness. The survey, which DNV GL fielded, ultimately provided the implementers with useful information on barriers to the adoption of the commercial financing program. It also told them which financing options the C&I customers preferred.

Reducing Evaluation Risk

As discussed earlier in this paper, there is a shared interest among program stakeholders in reducing evaluation risk. For its evaluation of the EU program portfolio, DNV GL identified three areas in which it could reduce evaluation risk without jeopardizing its independence as a third-party evaluator. These three areas included:

1. *Quarterly program tracking database reviews:* All of the residential energy-efficient measures and around half of the C&I measures in the EU portfolio have energy savings estimates which are deemed. If the program uses incorrect or outdated deemed savings values, this can result in the evaluators reducing program energy savings estimates.

To minimize this source of evaluation risk, in the first year of its evaluation, DNV GL conducted a program tracking database review prior to annual reporting. This review was comprehensive and allowed the implementation team to correct the database savings beforehand rather than creating an evaluation adjustment factor. Yet while this annual database review simplified some aspects of the evaluation, it also added complexity. Rather than simply calculating the right savings and reporting a verified total that was different from the program claimed total, the evaluation team engaged in a lengthy back-and-forth with the implementation team to justify or correct the savings reflected in the tracking database. This caused delays that put utility filing deadlines at risk and did not eliminate the risk of savings being changed at the end of the year.

To improve this process, DNV GL implemented a quarterly review of program tracking databases. A quarterly process reassured the implementation team that calculations were correct throughout the year. The added repetition and the spreading of corrections across the year also helped to reduce the time needed for the annual review.

2. *Adjusting the evaluation schedule:* In the early years of the evaluation of the EU program portfolio, in the fourth quarter of every program year the evaluators would survey participants from the first three quarters. This schedule had two disadvantages. First, when data collection happened in the fourth quarter of the year and analysis and reporting occurred in the first quarter of the following year, program implementers did not get the evaluation findings in time for their program planning efforts.

A second disadvantage was that the evaluators were not surveying customers who had participated in the last quarter of the program year. Often program participation spiked at the end of the implementation year. For example, due to their inherent complexity, many of the larger C&I projects did not get finalized until the end of the year. By only surveying participants from the first three quarters, the evaluation team had a “blind spot” for this fourth quarter program activity.

One fruit of the new, more collaborative relationship between program evaluators and implementers was the redesign of the evaluation schedule to avoid these problems. The new schedule benefitted the implementers by making evaluation results available in September when they could still inform program planning efforts. The new schedule also helped the evaluators by allowing them to evaluate program activities in the fourth quarter and by spreading evaluation activities out more across the year reducing pressure during the first quarter (just prior to utility filing deadlines).

3. *Establishing an engineering pre-review process:* Possibly the most important outcome of the more collaborative relationship between the EU program evaluators and implementers was to establish an engineering pre-review process. DNV GL set up a process where the program implementers could request an evaluation engineer review of documents for a proposed energy efficiency project prior to committing financial incentives to the project.

The pre-review was intended to be a high level “sanity check” on the project engineering calculations and assumptions. The evaluation engineer indicated where alternative calculations or assumptions might apply and pointed out where additional documentation was needed. The pre-review did not change the method that the evaluators would eventually use to verify the savings if the project was implemented, only the timing.

Both sides also found that discussing the calculations, assumptions and documentation prior to project completion helped to minimize some of the misunderstandings and contentiousness that can characterize post-evaluation conversations about engineering. The pre-review process also encouraged greater transparency from the program implementers than would have occurred in the post-evaluation review process. This was because the program has not yet invested significantly in the proposed project (both in terms of the incentives and an assurance to the customer about the amount of expected energy savings) and therefore there was less at stake in revealing the project’s limitations.

Since the introduction of the pre-review process, the evaluation team also noticed an improvement in the documentation on all projects, not just those projects targeted for pre-review. It is possible that the pre-review process had helped remind the program implementers about the importance of good project documentation.

However, the pre-review process did have challenges. One was the somewhat unpredictable and ad hoc nature of the pre-review requests. This put pressure on the evaluators to find, on short notice, engineers who are familiar with the project’s technology.

Even after the evaluation team found such qualified engineering resources, the pre-review process could present other challenges. For example, the customer whose project is undergoing pre-review might put pressure on the program to accelerate the project funding. This could, in turn, put pressure on the evaluation engineers to complete their pre-reviews more rapidly.

Conclusions

Our experience evaluating energy efficiency programs in Michigan over the past seven years has taught us the benefits of program evaluators working more closely with program implementers. Some of our high-level conclusions from this experience include:

1. *Program evaluators and implementers have a common interest in the success of the energy efficiency program.* An unsuccessful program can have many negative consequences including unhappy program participants, missed opportunities for energy savings, high levels of free ridership, the loss of performance incentives for utilities or implementation contractors, increased regulatory scrutiny, the replacement of the implementation contractor and the discontinuation of the program. While these negative consequences often weigh more heavily on program implementers, they can also create problems for program evaluators.
2. *Yet despite this common interest, relationships between program evaluators and implementers are often fraught with distrust and disappointment and can sometimes become*

antagonistic. One common reason for this is that program impact evaluators are often the bearers of bad news in terms of lower program realization rates. Program implementers also often complain that evaluation findings come too late to inform future program design.

3. *Lack of interaction between program evaluators and implementers can lead to lack of knowledge, empathy and trust between the two groups.* This lack of knowledge, empathy and trust can, in turn, lead to problems with program implementation and evaluation.
4. *Developing trust between program evaluators and implementers not only requires closer, more personal interactions between the two parties, but also a track record of openness and fair play between them.* For this reason trust between program evaluators and implementers usually cannot develop right away, but is built over time.
5. *Contract administrators play important roles in influencing the interactions between program evaluators and implementers.* They usually decide whether the program evaluation and implementation teams meet together or separately. They also serve as a filtering mechanism for any email or phone communications between evaluators and implementers. Reasons why contract administrators may want to keep evaluators and implementers apart include: 1) enhancing the appearance of the evaluators as being independent, third-party assessors of program performance; and 2) viewing their roles as necessary “translators” between two parties who speak different languages.
6. *It is helpful to have a contract administrator who can act as a “neutral mediator” between program evaluators and implementers.* As our Michigan case study shows, one advantage of having MCA as a contract manager is that they were a non-profit organization which did not benefit as directly from the program energy savings achievements as would be the case for a contract manager that was a utility with a performance-based incentive structure. This allowed MCA to better serve the role of “neutral mediator.”
7. *It is useful to have regulators involved in the program evaluator/implementer meetings:* One theoretical concern about closer interaction between program evaluators and implementers is that there might be a chance of cooption or collusion between the two parties. In our Michigan case study, the presence of regulator in the room helped to improve the transparency of the process. But besides their roles as contract administrators and regulatory “watchdogs,” the MPSC staff were a fertile source for ideas on ways to improve the design and delivery of the energy efficiency programs. Their involvement in the discussion also meant they had some ownership over any decisions to change program implementation or evaluation strategies based on interim program results.
8. *The positive experience with program evaluator/implementer collaboration in our Michigan evaluations led to some useful practices for minimizing evaluation risk and improving the usefulness of evaluation results.* These practices included:
 - a. *Establishing an engineering pre-review process:* In response to program implementer interest, DNV GL set up a process where the program implementers could request an evaluation engineer review of documents for a proposed energy efficiency project prior to the program committing financial incentives to the project. The pre-review was intended to be a high level “sanity check” on the project engineering calculations and assumptions. The evaluation engineer indicated where alternative calculations or assumptions might apply and pointed out where additional documentation was needed. Our multi-year experience with this engineering pre-review process has identified many positive aspects as well as a few challenges.

- b. *More timely evaluation findings:* For evaluation findings and recommendations to be “actionable,” they must be available when it is feasible for the program to take action. As our Michigan case study shows, we purposely adjusted the evaluation reporting schedule and increased the frequency of program tracking database reviews to try to reduce evaluation risk and improve the usefulness of our evaluation findings.