

Strategies and Policies for Improving Energy Efficiency Programs: Closing the Loop Between Evaluation and Implementation

*Edward Vine, California Institute for Energy and Environment/
Lawrence Berkeley National Laboratory*

ABSTRACT

Program implementers often use evaluation results to improve the performance of their programs, but, as described in this paper, this is not always the case. Based on a review of the literature, participation in workshops, and interviews with over 50 program implementers, evaluators, and regulators, the utilization of evaluation results is investigated by asking the following questions: (1) How are program evaluation results *used* by program implementers and other stakeholders? (2) How are program evaluation results *communicated* to program implementers and other stakeholders? (3) Are the needs of program implementers being met by program evaluation? (4) What is the role of the utility regulator in facilitating the use of program evaluation results? And (5) What other mechanisms can facilitate the use of program evaluation results? While there is some consensus on the answers to these questions, the type of interest in and use of evaluation varies by functional role (e.g., evaluator versus implementer), maturity of the energy efficiency market, institutional context (e.g., evaluation and implementation conducted inside the same organization, or evaluation and implementation conducted by separate entities), and by regulatory demands and evaluation interests.

Introduction

Evaluation plays a critical and strategic role in the promotion of energy efficiency. Conceptually, the typical energy efficiency program planning cycle starts with policy and strategic program planning, leading to the design of specific programs for particular markets, followed by the implementation of those programs, and concluding with the evaluation of those programs (e.g., impact and process evaluation). The results of the program evaluations then feed into the planning, design, and implementation of programs, and the cycle repeats itself. Thus, in this context, “closing the loop” refers to ensuring, at a minimum, that the results of program evaluation are provided to program planners and managers for consideration. There is also an expectation that some or all of the evaluation results are actually used by implementers¹ to improve the performance of their programs.

Encouraging the use of evaluation results by program managers and other stakeholders has become a more prominent and publicly visible topic as the budgets for energy efficiency program implementation and evaluation have increased. As a result, there is increased interest in the communication and use of evaluation results. As described below, the type of interest in and use of evaluation varies by functional role (evaluator versus implementer), maturity² of the

¹ Implementers refer to both utility program managers and staff who design the programs as well as third-party implementers who are in the field and do the actual implementation. Sometimes, information only goes to the former and not to the third-party implementers. Ideally, all parties should receive the evaluation results.

² Maturity refers to how many years a particular organization or region has been involved in implementing and evaluating energy efficiency programs, as well as the availability of funding for promoting energy efficiency. For example, more mature markets are represented by California, New York, and Massachusetts and newer markets are represented by Arizona, Kansas, and Missouri.

energy efficiency market, institutional context (e.g., evaluation and implementation conducted inside the same organization, or evaluation and implementation conducted by separate entities), and by regulatory demands and interests (e.g., evaluation as an accounting audit (report card) or evaluation as an information source for helping to transform markets by providing information on lessons learned or best programs).

We examined evaluation utilization based on a review of the literature, participation in workshops organized by the American Council for an Energy-Efficient Economy, the International Energy Program Evaluation Conference, and the Association of Energy Services Professionals, and interviews with over 50 program implementers, evaluators, and regulators in the U.S. and Canada. The interviewees were asked questions about the following five topics:³

1. Communication of program evaluation results: How and when are the results of program evaluation communicated to program implementers and other stakeholders?
2. Use of program evaluation results: How are program implementers and other stakeholders using program evaluation results? How does one know if the evaluation findings were accepted and acted upon?
3. Needs of program implementers: Are the needs of program implementers being addressed by evaluators? What can the evaluator or implementer do to ensure that these needs are being met?
4. Role of utility regulators: What is the role of the utility regulator? Should regulators mandate that implementers use the evaluation results? What other regulatory options are available to facilitate the use of evaluation results?
5. Mechanisms to facilitate the use of evaluation: What other mechanisms can facilitate the use of evaluation results by implementers and other stakeholders? What works and does not work?

This survey was informal and was not meant to capture all the activities of all the evaluation organizations that exist; hopefully, the information provided in this paper does reflect the general practice of evaluation of energy efficiency programs. The rest of the paper is structured around these topics.⁴

There are two organizational models where the implementation and evaluation of programs are administered and managed. In one model (the Integrated Roles Model), program implementation and evaluation are managed by the same entity (and conducted either by their own staff or by contractors). This entity could be a utility, state agency, or a regulatory commission. In a second model (the Separated Roles Model), program implementation is managed by one entity (e.g., a utility) and the impact evaluation is managed by another entity (e.g., regulatory commission). Each model presents special conditions that affect the use of evaluation results, as shown below.

³ It is important to note that several areas are linked: for example, poor communication arises due to the lack of trust between implementers and evaluators, so that evaluation results are not used; similarly, the lack of trust can occur due to poor communication, again resulting in the non-use of evaluation results.

⁴ Responses to the last category (mechanisms to facilitate the use of evaluation) have been integrated into the other topics and are not in a separate section.

Use of Program Evaluation Results

The energy efficiency program planning cycle described at the beginning of this paper is a simplistic model that assumes there is a linear flow of information going from the evaluator directly to the planners and implementers for changing program designs. In fact, the flow of information is often nonlinear and hard to predict, making it difficult to track the impact of evaluation results on program implementation and policy. Thus, the manner in which implementers use evaluation results should be viewed through the prism of a complicated reality. Multiple agendas are at work, and there are multiple forces bearing on potential changes in program design and policy. Evaluation results exert whatever influence they have within this rather complex context. As one interviewee noted, the results basically get fed into a mill, typically providing each stakeholder with various pieces of evidence to gain support for his or her particular agenda, and adding one more force to the existing array of forces that are pushing for or against possible changes in policy or program design. The result of this process is hard to predict, there may be a long delay before results get acted on, and the steps that ultimately get taken may bear little obvious relationship to the recommendations made by evaluators. Improving program design and operations is an ongoing process that involves learning from experience, identifying changing patterns of customer and trade ally response, and getting input from evaluators who systematically study the program. Thus, evaluation results may end up influencing program design and policy only at the margin, as the other forces in play may be much stronger. Similarly, we should not expect a correspondence between evaluation findings and recommendations on one hand, and subsequent changes in policy and program design and implementation on the other to be simple or clear cut. They seldom occur in this manner.

Given the above model, it is important to note that evaluation findings have been used to improve programs in multiple ways, such as developing or reviewing: (1) general program concepts, (2) strategies to reduce free ridership, (3) tracking systems to meet evaluation needs, (4) definitions and baselines, (5) deemed (stipulated) savings, and (6) algorithms and inputs for estimating savings. In fact, many interviewees noted that program implementers are taking more of an interest in evaluation results now than ever before, and they provided specific examples where evaluation results have been used by program and portfolio managers.

At the same time, while evaluation results sometimes get used, they may not always be used effectively or uniformly, most likely due to the environment described at the start of this section. While some evaluators did not know how implementers used the results from their studies, others thought the environment was getting better for utilizing evaluation results.

Interviewees noted that **the greatest use of evaluation results** occurs when: (1) implementers are very committed to changing and growing their programs and value any information they get: e.g., they change their forms, recruitment tactics, organizational structure, incentives, message, etc. to make the program more appealing to the end users and trade allies; (2) implementers have goals to reach and need to show progress; (3) implementer incentives are at stake; (4) evaluations are valued; (5) there is some level of management accountability in which the evaluation professional's opinions and findings are important; (6) utilities and/or their regulators require the implementation teams to address the recommendations in some way; (7) savings estimates or net-to-gross (NTG) ratios are assigned to the program measures; (8) implementers have a specific question or set of questions; and (9) evaluators provide specific recommendations that can readily be acted upon by implementers (e.g., changes in program design).

One recommended approach for increasing the use of evaluation results are “evaluability assessments” that implementers sometimes use as part of program design to see how evaluation information can be used in a program.⁵ As noted by Bronfman et al. (2008), evaluability assessment has a long history in the evaluation field, but has not been formally applied to energy efficiency programs. A program is able to be evaluated (evaluability) when: (1) program goals and priority information needs are well defined; (2) program goals are plausible; (3) relevant performance data can be obtained at reasonable cost; and (4) intended users of the evaluation results have agreed on how they will use the information. If these points are not met by the current program design, the program should have a timeline for working towards program redesign to meet these points. According to one evaluator who has conducted this type of assessment, the hardest part to complete is the fourth element above: how the information will be used.

Most interviewees agreed that the most effective approach for facilitating the use of evaluation results is a team approach. This is particularly true for the Integrated Roles Model. In the team approach model, evaluators work closely with the implementers to provide timely feedback, even before evaluation results are finalized, and managers discuss what changes are needed with the evaluation contractors. An important outcome of this approach and process is that implementers view evaluators as allies and resources for helping them to implement better programs, and evaluators better understand what is going on within the program and with the customers. Another outcome of this approach is that evaluators’ misunderstandings are corrected that otherwise would have resulted in inappropriate conclusions and findings.

As described earlier, implementers do not always use evaluation results, and those people who expect evaluation results to be used by implementers may be disappointed. Clearly, implementers, evaluators, and regulators contribute to this problem. For example, several interviewees noted many reasons why implementers are resistant to changing their programs based on evaluation results, such as (in no particular order): (1) there is a long lapse between implementing a program and a completed evaluation - program management often changes during a long lapse and, as a result, there is less ownership of evaluation results if they are not timely; (2) implementers are committed to their program design and structure, based on what works for them; (3) evaluators are held at arms’ length and are not involved in program design meetings and discussions; (4) implementers believe that evaluators should only evaluate existing programs and not make suggestions on how to improve programs because that would be interfering with program design; (5) implementers see evaluation as a nuisance and a distraction that takes them away from their daily activities; (6) implementers do not like to vary methodologies due to added costs and skill sets needed; (7) implementers have no financial incentive to change their program; (8) new evaluation methods may produce reduced savings estimates, thus negatively affecting their reported performance; (9) process evaluation findings are viewed mainly as assessments of the implementation team’s capabilities and tend to focus on problems; and (10) evaluation recommendations are sometimes not pragmatic or actionable, or seen as too costly or beyond the scope of the program.

On the other hand, several interviewees noted that evaluators contributed to the limited use of evaluation results by implementers. In a few states, the relationship between evaluators and implementers is hostile: according to one interviewee, evaluation contractors may not be

⁵ Initially developed by Wholey (1979), evaluability assessment seeks to gain information from important documents and input from stakeholders concerning the content and objectives of the program. Outcomes from this assessment include clear objectives, performance indicators, and options for program improvement.

listening to the program managers. Sometimes, the separation is so great that evaluation contractors only believe what they observe, even when it is wrong in the eyes of the program manager or the customer.

Finally, several interviewees noted that utility regulators influenced the type of evaluation results that get used by implementers (as well as by regulators). For example, in California, regulators are one of the primary audiences for impact evaluations, and the regulators focus on impact results that are used to support shareholder earnings claims that have been set up for program administrators. As a result, impact evaluations are generally designed primarily for auditing purposes (a “report card”) that limit the use of evaluation results by implementers⁶, and communications regarding impact evaluation results are likely to be contentious. Furthermore, when evaluations are designed for regulatory requirements, they are filled with specialized technical terminology (jargon), and there is often a need to “translate” the studies to an understandable level for educating implementers, which does not always occur.

For many organizations, there is no formal mechanism or process to make sure that programs consider evaluation recommendations, leading to possible adoption. In other cases, several organizations are pro-active in summarizing and tracking evaluation recommendations with program implementers to determine if the recommendations should be implemented and to document how the implementer has responded to each recommendation.

One useful approach for examining how evaluation results are used by implementers is to conduct an evaluation of the evaluations. This is exactly what the New York State Energy Research and Development Authority (NYSERDA) did when it reviewed over 200 actionable recommendations from evaluation studies using in-depth interviews and an email survey (Peters et al. 2007). In this review, the authors found that action was taken on 48% of the actionable recommendations that were suggested in the 2003 and 2004 program cycle, and that this percentage increased to 67% for recommendations suggested in the 2005 program cycle. The authors concluded that the evaluation capacity had increased for NYSERDA program and evaluation staff and that evaluation was increasingly seen as useful by implementers.

In summary, the use of evaluation results can be facilitated by conducting evaluability assessments, having implementers and evaluators work as a team, and tracking how evaluation recommendations are used by implementers.

Communication of Program Evaluation Results

The principle of getting evaluators involved very early in the program planning process and providing explicit feedback of results to program implementers is now fairly routine. Often, evaluation results are communicated to program implementers informally, during the evaluation process, through meetings, interim memos or reports, presentations, and discussions at monthly staff meetings and workshops. During this process, the evaluation team makes itself available to discuss evaluation results with program implementation staff on an as-needed basis. This is a “no surprises” strategy that reduces implementer’s concerns about the evaluation, gives them time to learn about results in “real time,” and gives them an opportunity to discuss the results. This does not mean that implementers determine the recommendations in evaluation reports. It is meant to

⁶ In California, the investor-owned utilities implement programs (with contractors or third-party implementers) and conduct process and market evaluations, while the regulatory commission manages the impact evaluations (the Divided Model).

avoid the disconnect that occurs when the evaluation recommendations are poorly communicated or are not provided in a timely fashion.

Final reports are often posted on websites, however, the final reports typically memorialize what has already been communicated, and, in some cases, the final report may no longer be relevant: e.g., changes recommended in the draft evaluation may have already been made by the time the final report is distributed. On the other hand, the final report can serve a useful purpose: e.g., some organizations include an appendix to the final report on how evaluators responded to the implementer's specific comments or concerns.

In addition to the final reports and formal presentations of results to program implementers, the results of the evaluations are presented at conferences and other meetings where implementers are in the audience and can share new ideas and concepts. Some organizations take further steps in communicating the results of evaluations to implementers and other stakeholders. In Oregon, the Energy Trust's program implementers work with the evaluation managers to write a staff response memo to an advisory board, which explains what was learned and what the organization is going to do about it. Program managers also sometimes sit in on meetings of the Board of Director's Evaluation Committee when evaluations of their programs are presented and discussed. At BC Hydro, the Evaluation Oversight Committee meets periodically during the year to review evaluation findings. Members of this committee represent various stakeholders in BC Hydro (e.g., generation, load forecasting, electricity planning, distribution planning, incentives and rates, Power Smart Marketing, Customer Care key account management) and from the BC Transmission to ensure that the evaluation findings are of value to the stakeholders.

In some areas of the country, the principle of getting evaluators involved early in the program planning process and providing explicit feedback of results to program implementers was more the exception rather than the rule. As a result, there is less likelihood that evaluation information is passed on to the implementer. There are at least two reasons for this situation. First, in some of the less mature markets, the same utility staff does both implementation and evaluation of programs (the Integrated Roles Model): on the surface, this seems like it would encourage communication between implementers and evaluators. However, program implementation staff are typically overburdened with program implementation of the new energy efficiency programs and often do not have the time to use their evaluation contractors to help implementers succeed. Second, programs that are primarily driven by regulatory requirements tend to be the least likely to have results (particularly, process evaluation results) passed on to the implementers. Under this regulatory culture, implementers are provided with the final results of the evaluations (e.g., kWh saved) but with little information on, for example, the best type of program delivery mechanism that resulted in those savings.

Where the communication process is not working, opportunities for improvement are available. First, evaluators can be proactive rather than reactive and participate in the implementer's planning sessions to bring their expertise to the upcoming program needs (the team approach). Thus, evaluators can be part of the assessment of program needs and program design, so that information and knowledge needs can be thoughtfully determined and provided in time to help with decisions. Second, evaluators can develop evaluation plans that communicate the goals, processes and expected products of the evaluation in a manner that is transparent and understood by program implementers.

Third, evaluators can write better evaluation reports. These reports should be easier to read and more understandable by non-evaluators. The recommendations in these reports should

be specific, clear, and actionable; in addition, if possible, the evaluator should report on the costs, implications, and effects of implementing their recommendations. Also, the reports should highlight what is good, and, when there are negative findings, the evaluator should provide constructive suggestions about how to improve the situation. The reports should not judge program managers and implementers negatively for every problem that crops up in a program. Fourth, the evaluation reports should be reviewed with those responsible for the programs to ensure that the presented findings are fair, accurate, and actionable, so that the final report and recommendations are seen as credible and useful by the implementers.

Finally, after the plans and reports are completed, evaluators should educate utility program implementation staff to understand how to interpret evaluation results (especially the reliability of the evaluation results and the uncertainty risks associated with these results) and how to communicate these results to implementers, including third-party implementers.

In summary, the communication of evaluation results can be facilitated by having implementers and evaluators work as a team, tracking how evaluation recommendations are used by implementers, and providing evaluation findings that are readable, fair, accurate, and actionable.

Needs of Program Implementers

In general, many interviewees thought that program implementers' needs were being addressed by process, impact, and market evaluations. Implementers are involved in attending project evaluation kickoff meetings, reviewing interim deliverables (e.g., project work plans, survey instruments, memos summarizing initial results, and draft reports), and maintaining open communications with evaluation staff throughout a project.

Nevertheless, there sometimes is a natural tension between evaluators and program implementers, and this model is not always viable (e.g., in the Separated Roles Model). Evaluators are under constant pressure to provide constructive feedback and support for program efforts. At the same time, evaluators are often torn between maintaining their distance from program implementers and retaining adequate rigor in methods and analysis, while being supportive of program objectives and program success. This tension is elevated when there is a firewall between evaluators and implementers,⁷ and, as a result, the needs of implementers may not be met. This tension is further increased when regulators and utilities are facing urgent needs – often for detailed impact results – and must primarily focus on those results related to meeting goals and incentive targets. As a result, there is less urgency about completing process and market types of evaluations that would provide useful information to implementers.

Several interviewees noted many reasons why implementers contributed to the problem of not having their needs met. First, implementers have trouble in communicating with the evaluators exactly what their program does, how the program operates, and what the intended results should or are hoped to be. Program theory and logic models help to articulate the program implementers' visions, but these tools are rarely used. Second, several evaluators felt that they were good at identifying issues and communicating them to implementers, but something gets lost in translation by implementers when the evaluators' recommendations are presented to those individuals and companies implementing in the field. Third, while evaluators make efforts to communicate their results to implementers in a way that is usable, implementers may become

⁷ Simply stated, the firewall prohibits implementation staff and contractors from evaluating programs and vice versa (e.g., in California).

defensive, stuck in their ways, and resistant to new information. The evaluators acknowledged that this was a learning experience for both implementers and evaluators, and they noted that when the same program people maintain a long-term relationship with the same evaluators over several years, then communication improves and the results more clearly match the needs of the program. In contrast to where personnel are regularly changing, the long-term relationships help to institutionalize knowledge and experience, leading to better understandings of different objectives. However, one evaluator warned that long-term relationships may lead to negative outcomes: the evaluation results may be tailored to what the evaluator perceives that the program wants to hear.

Several recommendations were suggested for implementers to help them to make sure that their needs are met. First, implementers should design a tracking system to quickly integrate data collection for closing the loop between implementers and evaluators. Second, program implementers (e.g., in utilities) should encourage their program implementation staff to emphasize the use of evaluation results to third-party implementers (e.g., contractors to utilities) and hold the latter accountable: for example, it could be in the implementer's contract that they must successfully address concerns raised by evaluation. Third, implementers should educate evaluators, so that evaluators understand that apparently excellent program design seldom matches a particular market perfectly. And fourth, implementers should have sufficient funding to track and collect program data and manage programs and support evaluations (meet with evaluators and talk about the program).

On the other hand, several interviewees noted that evaluators were partly responsible for not having implementers' needs met. First, evaluation results are often not provided in a timely manner. Second, evaluators are not always listening to implementers and regularly rely on evaluation test methodologies that don't measure the real impact of today's energy efficiency programs. For some implementers, these test methodologies are driven by utility cost-benefit tests that are 30 years old, are only relevant for rebate programs (and not information and education programs), and penalize the most creative and comprehensive programs using market intervention strategies. In particular, the energy savings from programs using behavioral (including operations and maintenance), community, education, and codes & standards initiatives are seen as too challenging to measure with most standard energy efficiency program measurement and evaluation approaches.

In response to some of these concerns, several interviewees noted that evaluation staff cannot be fully responsive to the needs of program implementation staff because they must give a higher priority to dealing with issues raised by regulators. They felt that they would need more internal evaluation staff, higher evaluation budgets, and more qualified evaluation consultants to be fully responsive to the needs identified by implementers. If they had these resources, they probably would spend more time (and money) focusing on targeted process evaluation questions and new technology assessments that would provide useful information to implementers.

Many recommendations were suggested for evaluators for helping to meet the needs of program implementers. In brief, evaluators need to: (1) make evaluation results more specific and directly relevant to the needs of the program; (2) use market research to determine what messages work, what motivates customers to act, what channels deliver best, and what partnerships can help programs thrive for evaluation purposes as well as for program design;⁸ (3) conduct attribution research, so that program implementers understand the effectiveness of their

⁸ For example, the Energy Trust of Oregon has conducted trade ally surveys and is beginning a residential awareness and perceptions survey.

program compared to other potential influences; (4) provide more real-time feedback on evaluation findings; (5) evaluate programs fairly and equitably, so that the implementer does not feel that his or her program has been unfairly selected; (6) conduct early evaluability assessments of programs; (7) present evaluation results orally (preferably in person), allowing implementers the opportunity to have their questions quickly cleared up; and (8) ask implementers directly what issues they would like to see evaluated to improve programs in the future.

To facilitate the process of meeting the needs of program implementers, the organization selecting evaluation consultants (or doing the evaluation themselves) should have a strong internal evaluation capability that can work with the program implementation staff and with upper management to identify the issues and to develop and design the evaluation activities. To achieve this goal, the organization should: (1) select an evaluation consultant who is familiar with the type of program to be evaluated and who demonstrates the ability to build positive relationships with program implementation staff; (2) have their evaluations scheduled so that findings can be incorporated into the organization's planning, budgeting, or program design cycle; and (3) have sufficient resources (time and money) for evaluators to address the needs of implementers (possibly including a contingency budget to address needs as they come up during the evaluation process).

It is difficult to do a good job on an impact evaluation and produce results less than fifteen months after the end of year. On the other hand, program managers desire faster feedback. Accordingly, several interviewees noted that implementers' needs for timely feedback regarding process issues and market response are not being satisfactorily met. Consequently, informal or interim feedback (e.g., informal meetings) is needed to make mid-course corrections or refinements. In response to this need, the Energy Trust of Oregon is developing a satisfaction survey on the internet that can meet evaluation needs to some degree and provide faster feedback. The Energy Trust is also experimenting with using in-house evaluation to provide simple first-line evaluations to provide program feedback. For small programs, this may be the only evaluation. For larger programs, more detailed evaluation studies may need to be performed by contractors. Also, if the results from the first evaluation show surprising results, a more detailed evaluation will be conducted before reaching conclusions. Furthermore, implementers might be able to meet their information needs more quickly on their own: e.g., residential lighting and appliance program implementers in New England routinely conduct mystery shopping and collect sales information and shelf surveys via their implementation contractors (vendors).

Finally, as noted previously, the most effective approach for meeting the needs of program implementers is a team approach where evaluation contractors and program implementers form a planning team (as in the Integrated Roles Model), so that the evaluation is focused on the needs of the program managers and on the documentation needs of the regulatory bodies. In the end, it is critical to have program implementers want their programs to be evaluated instead of thinking that evaluators exist just to kill programs. Finally, in order to really know if implementers' needs are being met, one may want to hire a third-party contractor to conduct an "evaluation of the evaluation," as was done by NYSERDA. An in-depth examination of the evaluation goals, processes, and outcomes can help identify areas where evaluation could better meet the needs of program implementers, so that modifications could be made.

In summary, the needs of implementers can be met more easily by conducting evaluability assessments, having implementers and evaluators work as a team, making evaluation results more specific and directly relevant to the needs of the program, providing more real-time

feedback to implementers, and providing sufficient resources for evaluators to address the needs of implementers.

The Role of the Utility Regulator

Regulators are often perceived as the primary stakeholder of evaluation activities among evaluators and implementers, and regulators themselves often see themselves as the ultimate users of evaluation information, as they use this information for a variety of activities (e.g., approving (or denying) utility program filings and evaluation budgets). Although regulators can mandate that utilities use evaluation information (e.g., load impact results on a forward-looking basis), regulators do not have the authority to dictate program specifics to utilities (e.g., mandate that specific evaluation recommendations be implemented for changing program design or implementation). However, regulators can play an active role in energy efficiency program evaluation by: (1) developing evaluation protocols and guidelines; (2) mandating that program evaluations be done (either by utilities or commission staff) and requiring evidence that implementers have reviewed, commented, and indicated that they are (or not) responding to the evaluation recommendations; (3) approving program implementation and evaluation budgets; and (4) reviewing evaluation results in contested case hearings.

Regulators can play an active role in formally encouraging the *use* of evaluation results by: (1) structuring incentives and disincentives so that utilities are motivated to pay attention to program evaluation results and take steps to encourage programs to grow and change: e.g., incentive mechanisms that incrementally reward better program results or give preferential treatment to programs that have been evaluated and that learn from past evaluation findings; (2) establishing specific forums (e.g., public hearings, workshops and forums) where evaluation information are shared with program implementers; (3) ensuring that evaluations are conducted in a more timely fashion, so that findings can be incorporated into a planning, budgeting, or program design cycle; (4) requiring utilities to support their energy savings calculations in program filings by referencing a regulator-sponsored database containing energy-related data based on past evaluation findings; (5) preparing a strategic evaluation plan that specifically indicates how the results from evaluation studies will be used by the regulator; (6) directing utilities to demonstrate that they have reviewed “best practices” information and to show how this information has been incorporated in each new cycle of program design; (7) ensuring that both process and impact evaluations occur; (8) providing budget flexibility to evaluators to have them work on key topics that arise from the implementer’s perspective; and (9) creating new performance metrics that goes beyond the standard Total Resource Cost (TRC) test and that would include such factors as non-energy impacts, customer satisfaction, and market effects that would be of great interest to implementers.

Regulators can also informally persuade utilities to use the results from evaluation studies: as an example, the New York Public Service Commission (NYPSC) places a high priority on evaluation, and while it does not have a formal policy regarding the use of evaluation results, NYPSC evaluation staff attend all meetings of the SBC Advisory Board when they review evaluation findings with NYSERDA evaluation staff and senior program management, and they review evaluation reports and emphasize to NYSERDA the importance of obtaining program recommendations from the evaluators and implementing the recommendations as appropriate.

Finally, regulatory staff handling evaluation issues can communicate evaluation findings directly with the regulatory staff handling program design issues. Regulators could follow up

with implementers to see whether recommendations are considered and either implemented or rejected. However, regulators often don't have the time, attention span, or access to all of the information needed to be sure that evaluation results are being used outside of the regulatory commission.

Most interviewees strongly felt that mandating an organization to use evaluation results was not a solution and that having well crafted regulatory goals, policies, and procedures regarding how savings are counted should make it in the best interest of implementers to incorporate the latest evaluation results on an ongoing basis. In fact, mandating the use of evaluation results may be counterproductive and may risk making worse mistakes than not using them.

Several interviewees also expressed concern about the regulatory environment itself, particularly in those states where regulators are primarily focused on impact evaluation and not on process evaluation. Furthermore, in those states, regulators are more focused on incentive programs rather than education, outreach, and marketing approaches that could vastly increase the effectiveness of energy efficiency programs. And in less mature regions, regulators may not understand the need for process evaluation, and, as a result, some of the programs may suffer.

There are at least two alternatives to mandating the use of evaluation results and regulatory intervention. First, some of the best examples of where feedback occurs are not as a result of mandatory actions but are due to the requirements developed by the funding organizations. For example, the Northwest Energy Efficiency Association, BPA, and the Energy Trust of Oregon require program managers to respond to evaluation recommendations in writing to management within 30 days of the preparation of the evaluation and to explain why (or why not) each recommendation was being addressed and what changes are being implemented, if any. And in the case of the Energy Trust, the Oregon PUC relies on the Energy Trust's Board of Directors for overseeing much of the Energy Trust's operations, including evaluation. The Board of Directors (one member is a PUC staff member) has a committee that reviews the drafts of evaluation reports, and explains the implications of the evaluation to the Board as a whole, with staff's assistance. The Board is very committed to seeing that the evaluation results are used, and sees that management is committed to doing so.

Second, the best approach to closing the loop is a collaborative evaluation process (the team approach) where the regulator, evaluator and implementer meet and talk about the questions that the evaluation should try to answer (as well as accurately defining the goals of an evaluation and the expected outcomes at the beginning of the project), and provide opportunities for them to participate in the evaluation process through interactive meetings and reviews of interim project deliverables over the course of the project. As indicated previously, the long-term relationships help to institutionalize knowledge and experience, leading to better understandings of different objectives and increased cooperation among all parties. Under this model, supporting the evaluation process and ensuring that the evaluation process is used and useful would be more important than mandating the use of evaluation results.

In summary, regulators can facilitate the use of evaluation results by having implementers, evaluators, and regulators work as a team, requiring that implementers have reviewed, commented, and indicated that they are (or not) responding to the evaluation recommendations, establishing forums for sharing evaluation information, requiring utilities to support their energy savings calculations based on findings from evaluation studies, and creating performance metrics that include factors such as non-energy impacts, customer satisfaction, and market effects.

Summary

The following key activities are needed for facilitating the use of evaluation results:

1. Implementers, evaluators, and regulators must work together as a team
2. Implementers must conduct evaluability assessments
3. Implementers and regulators must track how evaluation recommendations are used by implementers
4. Evaluators must provide evaluation findings that are readable, fair, accurate, and actionable - make evaluation results more specific and directly relevant to the needs of the program
5. Evaluators and regulators must provide more real-time feedback to implementers – for example, by establishing forums for sharing evaluation information
6. Regulators must:
 - a. Require that implementers have reviewed, commented, and indicated that they are (or not) responding to the evaluation recommendations
 - b. Provide sufficient resources for evaluators to address the needs of implementers
 - c. Require utilities to support their energy savings calculations based on findings from evaluation studies
 - d. Create performance metrics that go beyond direct energy savings and include factors such as non-energy impacts, customer satisfaction, and market effects.

We think many of these suggestions are applicable for both the Integrated Roles Model and the Separated Roles Model, although the team approach may not be viable under a Separated Roles Model.

We conclude with a note of caution: can an Integrated Roles Model “work” in a regulatory environment where program administrators are primarily judged by how well they attain specific energy savings goals, or how much money they can earn or lose? Is an institutional culture characterized by trust, transparency and a constructive approach to mistakes viable in this type of regulatory environment? Furthermore, the climate change issue has taken on more regulatory importance in many states. Will a carbon credit trading mechanism heighten the tension between implementers, evaluators, and regulators? If states and regions realize that energy efficiency is the first energy solution for reducing greenhouse gas emissions and that dramatic change is necessary to transform the market to become more energy efficient, then more partnerships will be needed, and these partnerships will have to work on a foundation of trust.

References

- Bronfman, B., S. Samiullah, A. West, J. Peters, N. Gandhi, and K. Cooney. 2008. “Integrating Evaluability Assessment into the Program Planning Implementation and Evaluation Process: Case Studies from Southern California Edison’s IDEEA Program Portfolio.” In *Proceedings of the 18th National Energy Services Conference*, Clearwater Beach, FL, January 28-30, 2008.
- Peters, J., S. Baggett, P. Gonzales, P. DeCotis, and B. Bronfman. 2007. “How Organizations Implement Evaluation Results.” In *Proceedings of the 2007 Energy Program Evaluation Conference*, 35-47. Chicago, IL: International Energy Program Evaluation Conference.
- Wholey, J. 1979. *Evaluation: Promise and Performance*. Washington, D.C.: The Urban Institute.