

What Will Improve Process and Market Evaluation?

Jane S. Peters, Research Into Action, Inc., Portland, OR

ABSTRACT

Process and market evaluation of energy efficiency programs is underrepresented in industry forums. Reviews of past IEPEC and ACEEE conference evaluation sessions, found less than 10% of the papers focused on process evaluation. Yet energy efficiency organizations typically allocate 25% to 30% of their evaluation budgets to process and market evaluations, suggesting industry conferences undervalue the impact and significance of process and market evaluations.

The benefits of process and market evaluations only occur if they are used. Identifying the types of uses for process and market evaluations, can provide clues to opportunities that increase their use. Standard evaluation utilization categories show that there are major contributions process and market evaluations offer to the energy efficiency industry.

At the same time, there are weaknesses in the current approach to process and market evaluation that require attention so that the potential value is achieved. Too much reliance on a limited palette of methods, limited application of social science theory in assessment and analysis, and a narrow focus as to what process and market evaluations should do limit the opportunity for process and market evaluation to inform program design.

Expanding the role of process and market evaluation has significant potential to take energy programs beyond the current focus on measured savings. Placing evaluation into a formative role, rather than a summative role, as process and market evaluation can do, will shift the focus from specific measure performance to an understanding of how to move energy efficiency into the mainstream.

Introduction

The research in this paper was part of a larger study commissioned by the California Public Utilities Commission (CPUC) and managed by the California Institute for Energy Efficiency (CIEE). In 2007 and 2008, we conducted interviews with 58 evaluators and program implementers, and reviewed over 100 documents either recommended by the interviewees or identified in a review of conference proceedings (Peters & McRae 2009).

The goal of the research clarified current use of process and market evaluations for energy efficiency programs and identified opportunities to improve and enhance process evaluation. Through this process, we hoped to bring a late 20th Century investigative approach into the 21st Century.

Description of Process and Market Evaluation

In the mid-1970s, policymakers began to ask questions about the effectiveness of efforts to increase energy conservation in homes, businesses, and industry. Since that time, evaluation for energy efficiency programs has been increasingly formalized, with protocols and guidelines developed to address issues specific to energy efficiency.¹

¹ The history of this process for California is discussed in the *California Evaluation Framework* (TecMarket Works Team 2004).

Process evaluation and *market evaluation* refer to specific types of evaluation activities. Market evaluations are typically conducted for programs that target market change, while process evaluations are typically conducted for any type of program, including non-energy efficiency programs. Market evaluations are less common outside of energy efficiency, as few public policy efforts so directly seek to change market behavior.²

Definitions for *process evaluation* point to several key features of process and market evaluations – specifically, that they are systematic investigations for the purpose of improving program implementation (Hirst & Reed 1991; Spinney, Peters & O’Rourke 1992; TecMarket Works Team 2006).

These definitions are consistent with definitions within the broad field of evaluation. For example, Davidson, a leading writer in program evaluation methods, says: “*When we evaluate the ‘process’ of a program or policy, we are talking about taking a critical look at the quality or value of everything about the program (what it is and does) except outcomes and costs.*” (Davidson 2005, 56; emphasis in original)

The focus of process evaluation is broad, including market evaluation, implementation assessment, needs assessments, and assessment or evaluative activity that does not address outcomes and costs. Thus, the conference of the American Evaluation Association is largely focused on what energy efficiency evaluators term *process evaluations*, across a variety of subject areas: education, health care, public policy, etc.

Such is not the case for energy efficiency program evaluation conferences, where the focus is largely on impact evaluation results and methods. Confirmation of this can be seen in reviewing the paper and panel sessions for the 1987, 1995, 2007 and 2009 International Energy Program Evaluation Conferences (IEPEC). The number of sessions with *process evaluation* in the title and the number of papers addressing process evaluation are meager. In 1987, two (6%) of the 32 sessions at the conference included *process evaluation* in a session title. At the 1995 conference, none of the 40 sessions included *process evaluation* in a session title, though *process evaluation* was included in five paper titles. The 2007 IEPEC agenda shows increased interest in process evaluation, with 5% of the 39 sessions focused on process issues; though no session titles or paper titles referenced *process*, there were 25 papers with process evaluation results reported. In 2009, 4 of the 36 sessions addressed process issues (two with *process* in the title) and 18 papers reported on process evaluation results.

Common Methods and Data Sources

The methods commonly used for process and market evaluation include: project scoping activities; data collection activities; in-depth interviews with individuals or groups; survey research by phone, electronic, or mail; reviews of data-tracking systems, program documents, and materials; direct observation; focus groups; and data analysis using quantitative and qualitative techniques.

Common data sources include: the utility or program administrators; third-party program implementation staff; stakeholders; participants; nonparticipants; and participating and nonparticipating market actors, including trade allies, vendors, design and construction professionals, retailers, distributors, and manufacturers.

At its best, process and market evaluation assesses the entire scope of the program to provide feedback on all its facets. When this does not happen, it is often a reflection of the lack of resources allocated to evaluation efforts, or to the process evaluation specifically. This lack of resources and focus was apparent for the New York State Energy Research and Development Authority’s (NYSERDA) first evaluation cycle (Peters et al. 2007) and in the California 2002-2003 program cycle (Hall et al. 2005). In both cases, process

² A search on the American Evaluation Association website for *market evaluation* found 53 results, including conference presentations, resumes, and company profiles; a search for *social marketing* found 15 results, most of which were for health communications, or energy efficiency studies.

evaluations were not funded and even impact evaluations were funded at such a low level as to not provide valid and reliable results.

Benefits of Process and Market Evaluation

A key concern with evaluation is whether the results are used: termed *evaluation utilization*. Evaluation utilization refers to the degree to which the findings of evaluation influence the programs, processes, or policies investigated, or affect the organization in which these occur (Dahler-Larsen 1998; Forss, Rebien & Carlsson 2002; Grasso 2003; Hodges & Hernandez 1999; Mark 2004; Valovirta 2002; Weiss 1998).

Owen (1999) and Weiss and Bucuavalas (1980) define four primary types of evaluation use:

- **Instrumental Use** – evaluation directly influences decision-making or the program itself
- **Conceptual Use** – evaluation influences thinking about the program, even if it does not lead to decision-making
- **Legitimative Use** – evaluation is used to justify decisions made and respond to concerns of policymakers for continued support
- **Process Use** – evaluation is of cognitive and behavioral changes resulting from the user's involvement in the evaluation process (enhanced skills in evaluation, routine incorporation of evaluation in the organization's processes, and program improvement – also termed *capacity building*)

Process and market evaluations, if they are to benefit the programs examined, need to be useful and used. There are three ways that evaluations benefit programs: as a management tool (i.e., instrumental, conceptual, and process use); to improve cost-effectiveness (i.e., instrumental and legitimative use); and to understand customer and market perceptions (i.e., instrumental and conceptual use).

Weaknesses of Process and Market Evaluation

Process and market evaluation has many benefits, but is also subject to some weaknesses. Any evaluation is just a snapshot, a single point-in-time when the evaluators observe program activity. Process and market evaluation particularly needs to be careful.

One of the contacts with a program implementation focus noted: *“You need lots of persistence and long-term presence in the market to get at who you need to talk to. I have reviewed some process evaluations [that did not reflect this] and wondered, ‘Would a program designer be able to learn enough from this process evaluation to be able to grasp issues?’ The research just didn’t go deep enough.”* (Peters & McRae 2009, 12)

Another weakness of process evaluation is what one contact referred to as *“no more than gossip.”* In such cases, the evaluator focuses on “who said what” or focuses on whether participants are satisfied with the program. Focusing solely on opinions and satisfaction may meet a regulatory requirement, but may not provide insight to guide program improvement.

Process evaluations must go beyond satisfaction and gossip and internal processes of the administrator's organization to understand the program. They also should have sufficient budget and the time required for the evaluators to understand the program well.

Improving Process and Market Evaluation

Process and market evaluations provide assessments of the program as designed, intended, and stated in the goals and objectives; they are not primarily an exploration of the issues or concerns of immediate

interest to the program staff, but rather a *specific intent*. Toward that end, there are three areas where we can improve our evaluations: improve our cultural competence, improve our methods, and connect with implementers.

Improving Cultural Competence

While process and market evaluation may be appropriate for virtually all programs, the adage “one size does *not* fit all” applies. Conducting a process evaluation for a program in China required sensitivity to the culture of program implementers (including cultural styles of learning and self-assessment), as well as the culture of the Western program funders (Peters, Nadel & Guijin 2005). There are distinctions among residential, commercial, and industrial programs, which similarly require cultural competence on the part of the evaluation team seeking information from these customer groups, and segments within these groups. A market transformation program may have a different cultural basis than a resource acquisition program, and programs operated in Illinois have a different regulatory culture from those in California.

Improving process and market evaluation begins with the evaluator improving cultural competence relative to: the program, the organization operating the program, the target market for the program, and the regulatory and legislative environment in which the program operates. High cultural competence enables the evaluator to be sensitive to the issues, to be thoughtful about the recommendations, and to be considerate of the market and market actors.

Improving Methods

The most common primary data collection methods used for process and market evaluation are surveys and in-depth interviews. These data methods result in a combination of qualitative and quantitative data that must be analyzed and interpreted in order to create an understanding of the program experience from multiple points-of-view. There are also a variety of other data collection techniques and a variety of ways of handling these types of data.

Surveys. Survey research is a core methodological element for evaluating energy efficiency programs. Most evaluators use surveys as a component of their investigations. Surveys of participants, partial participants (those who have had contact with the program, but exited the program without completing a project), and nonparticipants (those who, when screened, are found to qualify, but have not had contact with the program) are key sources for assessing how the program is operating in the field.

Several contacts, however, noted that surveys can comprise too much of the evaluation efforts: “*We rely too much on surveys. We need to do more open-ended interviews and process analysis. Not enough attention is paid to existing records and data. If survey data could be joined with other data sources, it might be more powerful.*” (Peters & McRae 2009, 47)

Clearly, surveys are not used as well as they might be. Process evaluators need to look for more than just satisfaction or dissatisfaction when they develop survey questions. One problem noted by several contacts is that satisfaction levels with programs are typically quite high. As Hall et al. (2005) noted in their review of process evaluations for the 2002-2003 California program cycle, satisfaction questions tended to be poor, while in-depth questions on operations were better. They recommended that satisfaction questions be focused on program improvement and be designed to enable comparisons across programs.

An approach that may be more effective than satisfaction questions is to examine *service quality* (Rust & Oliver 1994). Service quality focuses on whether the program experience met or exceeded expectations and examines the value obtained from the program. Questions can also identify what

expectations were held by participants, thus revealing opportunities for improvement and enhancement to the program.

The use of surveys primarily to gauge response specific to a program year is important, but this approach misses an important opportunity to use surveys to assess program progress over time (Barata and Anderson 2007). The key step that needs to be taken, as Barata and Anderson note, is that the evaluator establishes a benchmark early on, possibly before program implementation begins. The benchmark provides a basis for assessing change and for gaining insights into how to improve the program at the outset of the implementation period.

Finally, surveys have a potential for bias when evaluations rely on participant and partial participant points-of-view. Unless nonparticipants are sampled as well, the surveys will likely portray an unrealistic perception of the market reaction to the program. At a minimum, evaluators should be clear as to the biases, rather than assume that the participants and partial participants represent the market.

In-Depth Interviews. When done well, in-depth interviews are very effective; when poorly done, they can be “little more than gossip,” as one contact put it (Peters & McRae 2009, 48). Bensch and Pigg’s (2002) review of process evaluations and market research found that: “*[Well designed] in-depth interviews – by themselves or in conjunction with other qualitative data... provide a window to the decision-making process and the world of the market actors that energy-efficiency programs have been trying to influence – particularly when [the interviews] address the perspectives of multiple market actors in the context of a specific product or service.*”

In-depth interviews require structured interview guides. Nicol and Shepp (2005) noted how they failed to have a detailed guide and tried to complete the interviews in a short time, missing many opportunities to get more detailed information. The guides need to be structured and, as one of the contacts pointed out, they need to be developed “*in conjunction with program procedures and data.*” (Peters & McRae 2009, 48)

It is also important to allocate sufficient time. Nicol and Shepp (2005) noted that they should have allocated at least 30 minutes for the interviews. In-depth interviews often take an hour or more. While everyone is very busy, the evaluators need sufficient time with the staff and key market actors to fully understand the program.

In addition to allocating sufficient time and having a structured interview guide, it is also optimal to have more than one person involved in conducting the interview (Peters, Nadel & Guijin 2005). When two people conduct interviews, one researcher can focus on the respondent while the other takes notes and records the interview. After the interview, the two can discuss the interview, identify opportunities to improve future interviews, and begin to generate hypotheses for further investigation during forthcoming interviews and in the analysis process. While these steps can and should be conducted by the individual interviewer working alone, two interviewers tend to improve the reflection and analysis process; however, the costs will be higher.

Focus Groups. Focus groups – where the evaluator moderates a discussion with a group of contacts – that center on a specific or narrow aspect of the program are not used as frequently as surveys or in-depth interviews, but can be useful for data collection on program features.

Focus groups are perceived as being inexpensive, since one can talk to a dozen people in a two-hour period, instead of the over 10 or 12 hours that would be required for in-depth interviews. However, this is largely a misconception, as focus groups can cost more than 10 to 12 hours of in-depth interviews conducted by phone at the respondents’ convenience. The primary benefit of focus groups is interviewing several people at once about a common topic and witnessing the interaction between respondents. Focus groups are

not a replacement for systematic survey research or in-depth interviews, since too few people can be included to represent the population and there is so little time with each respondent.

Focus groups are effective when the goal is to isolate program features, including marketing messages, team processes with account representatives, or reasons why a form or a program participation letter may not work. They can also be an effective way to test terminology and focus for a survey, to gain insights into why respondents provided certain responses, or to obtain reactions to conclusions or findings. When focused on a narrow topic, such as those noted, they can be very effective tools to augment and enhance other data collection efforts.

Other Methods. This section addresses some of the other methods that can be used for process evaluations. As more than one contact put it: *“We need to rely on information other than surveys... look for ways to verify what they are saying.”* (Peters & McRae 2009, 50) There is a variety of additional data collection methods, as well as analytical methods that can be applied and have been applied in process and market evaluations.

Onsite Data Collection – collecting data onsite rather than by phone, mail, or the Internet, “increases insight” as one contact put it (Peters & McRae 2009, 50). Onsite data collection takes a variety of forms. Many of the same risks of using focus groups apply to onsite data collection, due to the usually small sample sizes. Care must be taken to draw samples appropriately and to draw conclusions consistent with the data.

The first type of onsite data collection is to do in-depth interviews in-person. This sometimes is beyond the resources available for the evaluation, but it clearly is optimum. Contacts spend more time with the interviewer and, because they are on their “home turf,” those interviewed feel more comfortable and more willing to discuss their point-of-view.

Onsite data collection can also include investigations of installations and engineering process evaluations as described by Kallock (2005). These efforts enable the evaluation to observe early in the program history how installation and procedures are being implemented in the field at a sample of sites.

Onsite data collection also includes “ride-along observations” with the implementation, sales/circuit-rider, or energy assessment teams. This provides an opportunity to see what happens during customer interactions. However, one observation cannot be generalized to all crews and, at most, should be used to inform other data and to help shape survey questions, unless the sampling is sufficient to permit generalization.

The final type of onsite collection gets into the realm of market research “mystery shopping” and “stocking counts” etc. These approaches have professional shoppers go to stores to shop for products, make calls for services, assess what is on the retail floor, or assess how a wholesaler handles requests for efficient equipment. These approaches, like ride-along observations, provide a more realistic assessment of the purchase experience.

Verify Market Information – There are many ways to verify market information. Two contacts discussed how they had used Geographic Information Systems (GIS) data to locate participants in the utility service territory (Peters & McRae 2009, 51). One contact analyzed GIS data for a low-income program, which enabled the program to target census tracts with door-to-door appeals. An example of this is a study that examined participants in California programs to determine if there were geographic areas with high and low participation rates (Wirtshafter and Sammiullah 2005).

Another example of market information verification was a market simulation of market conditions by Peters and Tannenbaum (1997). The project randomly assigned survey respondents to receive marketing materials from one of the five potential delivery agents and included descriptions of the selected delivery agents’ roles in the program. This provided the framework for testing reactions to the program concept, as well as the different delivery agents.

Data Collection and Theoretical Approaches. The following analytical or theoretical approaches can result in unexpected or new findings and demonstrate how evaluators who veer from strategies typically used, may provide new insights to program effectiveness.

Real Time Data Collection – data are collected during program implementation, analysis proceeds at staged intervals, and short reports are provided to permit mid-course corrections in the program. While there are many examples of this approach, it is still limited, and requires coordination with the program team and the evaluation team at the outset of the program design process to ensure that data collection points are integrated. An example is the *WashWise* program by the Northwest Energy Efficiency Alliance (NEEA). The first three reports were completed at six-month intervals. Subsequent reports were increasingly less frequent as market transformation effects continued and NEEA determined less intervention was required.³

Social Network Analysis – examines the network of relationships between people who are in communication with each other, and measures the strengths and weaknesses within the network. With the advent of electronic data and advanced computing capability, one can explore electronic information flow (emails or Internet traffic) whenever the data analyst can obtain traffic for actors in a program. For example, for the process evaluation of California's 2006-2008 Statewide Marketing and Outreach Program, the authors used social network analysis to assess the communication patterns among the stakeholders and team members based on the email traffic between the various members (Opinion Dynamics Corporation 2008).

Concept Mapping – developed by William Trochim (Kane & Trochim 2007) to collectively bring groups to a common understanding of a problem or issue. It is essentially a consensus-building tool that uses multi-attribute analysis to find commonalities across statements and ratings made by a group of people regarding a concept or idea. Lee (2007) used concept mapping to develop a program theory and logic model.

Mental Models – a theory-grounded strategy for exploring how and whether change has occurred as a result of program interventions. Respondents (program staff, as well as program participants and dropouts) are asked about their views or their mental model at various times in the intervention process. Wilson, McDaniels, and Bennett (2006) used mental models to explore the effects of trying to change major commercial developers' approaches to the design of projects in a large city.

Cultural Models – another theory-grounded strategy for exploring program influence. A cultural model perspective suggests that the evaluator understand the cultural context of the program manager and program participants in order to gain greater understanding of how the program is affecting change. Janda (1994) recommended that a cultural analysis be taken in analyzing response to a program that focused on professional designers and new commercial construction, largely because designers are trained to have a very strong culture of practice.

Pilots and Experiments. Before programs are finalized, they can be offered in pilots and experiments. While program tests are infrequent, experience suggests that coupling process and market evaluation activities will both improve the ability of the program to meet its ultimate objectives and warn program administrators of problems before program launch.

The Niagara Mohawk experiment of marketing alone or with different incentive options (Mast & Ignelzie 1996) demonstrated an approach for experimenting and testing concepts. As Sullivan (2009) points out, these types of experimental designs can be used more often to test different delivery options, for different groups of customers, in different conditions.

The High Efficiency Laundry Metering and Market Analysis (THELMA) project was an energy-efficiency venture that sought to address a variety of market and technology issues upfront (Electric Power Research Institute 1998). Designed to investigate the benefits and costs of energy-efficient clothes washing

³ The various reports can be accessed from the NEEA ENERGY STAR® Consumer Products Reports website: <http://www.nwalliance.org/research/evaluationreports.aspx>.

machines, the project included detailed impact evaluation efforts, along with market and process assessment components. The researchers conducted in-depth interviews with customers to understand their experiences with the energy-efficient washing machines and the participating manufacturers to improve the machines. When the utilities offered rebates to support the purchase of these new washers, the machines were technological winners (Shel Feldman Management Consulting, Research Into Action & Xenergy 2001).

Process evaluators with engineering skills can examine the effectiveness of new ideas and technologies. Gandhi et al. (2007) describes a series of small-scale programs using innovative technologies that were funded through the Southern California Edison Company's (SCE) Innovative Designs for Energy Efficiency Activities (IDEAA) program in 2004 and 2005. Several of the products were found to be "inappropriate for the application without further development and the selected products did not offer the claimed benefits under all operating conditions." Understanding the types of problems identified in the process evaluations, SCE was able to prepare for the impact evaluation.

Gaining the Implementers' Attention

Another means to improve process and market evaluations is to get the implementers' attention – gain their confidence so they seek and support more research to improve the efficiency and effectiveness of their programs. Three opportunities include: evaluability assessment; theory and logic modeling; and improved recommendations.

Evaluability Assessment. Evaluability assessments are common in the evaluation community. Their purpose is to investigate at an early stage of implementation whether a program is collecting the right data, and whether the goals and objectives are sufficiently defined so that an evaluation can be conducted and draw meaningful conclusions (Bronfman et al. 2008). One contact, who had just begun to use evaluability assessments noted: *"Managers get information about how well they are managing the program, including the data collection needs. This effort often discovers new things on the program. [It] helps you scope the evaluation, quickly assess, [and] avoid wasting evaluation resources."*

Theory and Logic Models. Ideally theory and logic models⁴ would be the province of program managers; however, many evaluators have found that program managers are uncomfortable and unfamiliar with how to develop them. It often falls to the process and market evaluation teams to develop the program theory and logic models for the program managers to review and on which to comment. Program theory and logic models are very useful for delineating key issues for the evaluations to explore and, when used in conjunction with evaluability assessments, can alert both program managers and evaluators to important issues.

Useful Recommendations. Vine (2008) examined how evaluation can be used to improve program implementation. He also noted that evaluations of evaluations can be useful in improving the evaluation process itself. A process evaluation of the NYSERDA evaluation process, Peters et al. (2007) found that for recommendations to be received and considered, they needed to be actionable and realistic. As one of the contacts described it: *"[The evaluator needs to] understand the program implementer's constraints, understand the timing. If evaluation is just a regulatory requirement, it feels more like an audit. If there is*

⁴ Theory and logic models are the combination of a narrative theory for how a program is expected to achieve outcomes and the graphical representation of that theory. A theory and logic model does not guarantee that the theory is logical or that the outcomes will happen as depicted in the theory and logic model; rather, it is a draft of the intention. Theory and logic models typically include the program activities, the immediate outputs, and the short-, medium-, and long-term outcomes for the program (Knowlton & Phillips 2009; United Way, 1996).

lots of motivation to find something out, make sure you can target that – be useful. There is a need to effectively interact with and educate the implementation people about process evaluation. You want to make sure to get them on board, but also make sure everyone knows the ‘rules’: who the audience is, what the implementer’s role is.”

Evaluations can feel like a report card, but if the recommendations are actionable, it is clear who can take the action, and if the action is within the scope of the program manager and their management, then the recommendations are more likely to be implemented.

Opportunities and Challenges

Bensch and Pigg’s (2002) review of 311 evaluation reports noted that many of the reports tended to be repetitive of one another – process evaluations of similar programs across multiple program implementers. Bensch and Pigg also found a lack of depth in understanding decision-making processes.

As Kunkle and Lutzenhiser (1999) point out, there has been a great deal of effort spent on estimating free-riders, spillover, self-selection bias, and net and gross savings. This effort has kept evaluators from looking more deeply into the decision-making and the theoretical assumptions underlying the programs. They exhort the industry to: “[See the role of evaluation] in the context of theory that explains how we expect a particular market intervention to be successful. Evaluation must test the assumptions behind the underlying theory for a particular market intervention by determining if the expected market changes occur, identifying the motivating factors for the market actors related to these changes, and incorporating the costs and benefits as seen through the eyes of the market actors.”

The opportunities to improve process and market evaluation and use them as the vehicles for testing assumptions about how the market works are the most important avenues of change that we recommend for process and market evaluation. These research efforts should be examining:

- What is motivating the different market actors?
- What are the most effective delivery mechanisms?
- How are decisions made in organizations and in households about energy, and how do these decision-making processes differ depending on the types of energy efficiency action?
- What typology can we develop for the different actions that could accelerate adoption and implementation?

Arms-length evaluation. Some of the contacts (Peters & McRae 2009, 56) noted that process and market evaluators need to be “close to the implementers” in order to have any influence on the managers’ decisions or to fully understand the program and to gain “context rather than just a point-of-view.” Other contacts warned of the “need to be careful,” recognizing that when the evaluators get too close to the program implementers, they may fail to make recommendations or even fail to see what is occurring in the program. For example, when a program manager provides a list for interviews, is it a complete list of the population or is it a selected sample?

As O’Drain and Peters (1993) describe, there is a process that occurs when evaluators begin data collection that leads them to be sympathetic to the program manager (the first point of in-depth contact) and later to other parties to the program. Only through conducting a full assessment of the program from multiple points-of-view is it possible for the evaluator to gain an open and objective point-of-view.

Our perspective is that this tension is actually good for process and market evaluation, if the evaluator takes it seriously. A concern that Hall et al. (2005) expressed regarding the process evaluations for the California 2002-2003 program cycle was their lack of comprehensive assessments of the programs. If evaluators keep in mind the need to expand their data collection to multiple parties, and not focus on just the program staff and participants, then evaluations will be more likely to be objective. In addition to collecting

data from all sources pertinent to the assessment, it is also important to provide an honest assessment – as one contact described (Peters & McRae 2009, 57), “*At a minimum, the draft report should be as honest as possible.*” Ultimately, the more honest and direct the draft report is, the more likely that the evaluation will be effective.

Supporting these suggestions are the American Evaluation Association’s (AEA) *Guiding Principles for Evaluation* (2003) and the Joint Committee on Standards for Educational Evaluation’s *Program Evaluation Standards* (1994). These principles and standards are important to adopt and practice. Additionally, regulators and clients can use them to assess evaluations and evaluators, as long as there are sufficient resources allocated to do this degree of investigation.

Market Assessment or Market and Process Evaluation? Some of the contacts (Peters & McRae 2009) expressed concern that market research not merge with market evaluation to the point where marketing research is no longer done or, conversely, that market evaluations are not done. There is a difference between market evaluation and market research. In large part, marketing and market research should be conducted while programs are being designed, and should directly influence program design and development. Market evaluation, in contrast, is evaluative.

Market and process evaluations have clearly been a major source of information about the market and about consumer and business response to energy efficiency; however, this does not replace market research. Market and process evaluation provide ongoing feedback about the market in the context of the program, and should be used to evaluate how well goals and objectives are being achieved. Market and process evaluations should take the market context into account relative to the program..

Coordination with Impact Evaluations. A number of contacts stated that process evaluation should be coordinated with impact evaluations. The purpose would be to enable impact evaluations to be more effective and better able to interpret findings. Yet perhaps, as one contact noted (Peters & McRae 2009), the key difference should not be between impact and process evaluation, but between summative and formative evaluations.

This notion of thinking in terms of formative and summative is important. Impact evaluations are typically summative, completed after sufficient data on energy use are available to ascertain if there has been a change as a result of the program. This inevitably means that the summative impact evaluation will be completed 16 to 20 months after a program year is over. Many process evaluations are conducted as part of comprehensive evaluations – including impact and process evaluations. Yet these efforts are primarily summative process evaluations, as they do not provide information to directly improve the program while it is being implemented.

Yet, process and market evaluation are most effective if they happen early and often in the program implementation cycle, when they have the most impact. There are formative evaluation opportunities for engineers and economists as well. As discussed above, Gandhi et al. (2007) did formative studies for new technologies. This might be called an engineering process evaluation and is a mechanism to explore impact issues in a formative manner.

Designing an evaluation to be either formative or summative, rather than process or impact, would potentially lead to the question of what formative evaluation research is needed in year one, year two, year three, etc., and, conversely, what summative evaluation research is needed to understand what happened in year one, year two, etc. Such a consideration would lead to more process and market evaluations, with engineering and economic components conducted early in program implementation cycles to provide formative input to the programs. In subsequent program cycle years, there would be impact evaluations with some process and market components providing a summative assessment of the program.

The other key issue regarding coordination of impact with process and market evaluation concerns timing (Vine 2008), getting evaluations into the field early enough to be useful. Rethinking evaluation in terms of what are the formative evaluation research needs and what are the summative evaluation research needs would enable evaluation designs to more effectively match the needs of program designers, planners, implementers, regulators, and stakeholders.

Conclusions and Recommendations

There have been many evaluations that have sought to understand the *how and why*, but the key recommendation for improving process and market evaluation is to expand efforts to explore the how and why with increased emphasis on formative evaluation.

At this point, the methodological tool bag for process and market evaluation is full of underutilized options. Evaluators currently rely too much on surveys. In addition, while in-depth interviews are common, they are rarely conducted in-person due to increased costs. Focus groups tend to be misused as a cheaper alternative to surveys and in-depth interviews, instead of being used for their explicit purpose – to focus on narrow issues that need illumination within the context of other data collection approaches.

The underutilized options are numerous and include: on-site observation; engineering process evaluations; GIS analysis; market simulations; social network analysis; concept mapping; mental or cultural model analyses; and the use of pilot programs and experiments to test new ideas and explore how effective they might be, and what is required to achieve maximum effectiveness. Additionally, we highly recommend the use of evaluability assessment and logic and theory modeling as tools to engage evaluators and program implementers early in program processes to ensure evaluations will be effective.

Two activities can facilitate an expansion of process evaluators' ability to look deeply. The first is to encourage process and market evaluators to draw upon their social science backgrounds and think about how social science theory explains what is, and what is not, happening in the programs that they evaluate. Each evaluation should explicitly explore whether the theory hypothesized in a program theory and logic model is working as intended and, if not, what might be another explanation. The second key activity is to encourage a greater connection to the broad program evaluation community. The American Evaluation Association sponsors an Evaluators Institute⁵ to train evaluators in various techniques of evaluation and provides certification to those who complete the full training series. The guidelines for evaluation practice, as noted earlier, can help guide policymakers, program implementers, and evaluators to assess program evaluation and structure evaluation studies to best meet their needs.

Beyond looking more deeply at programs, we believe there is a need at this point in the energy program cycle to consider whether the division of resources between process and market evaluation, as compared to impact evaluation, is truly meeting the needs of the market. Process and market evaluation remain underutilized and relatively invisible compared to impact evaluation. Impact evaluation typically occurs too late in the program cycle to have much effect on the program, as it is largely summative in nature.

Evaluations that can actually improve programs need to be executed early in the program implementation process to provide feedback.

In summary, we recommend the following:

- Move to a framework for evaluation that promotes *formative* evaluation during the program implementation cycle, and *summative* evaluation to explore program effects and impacts after reasonable periods of time to obtain results, rather than combining process or impact evaluation.

⁵ See: <http://www.eval.org/SummerInstitute08/08SIhome.asp>.

- Increase the use of process and market evaluation to test assumptions and market conditions early in program implementation.
- Expand the range of methods used and resources available for process and market evaluation research to include on-site observation, engineering process evaluations, GIS analysis, market simulations, social network analysis, concept mapping, mental or cultural model analyses, and the use of pilot programs and experiments.
- Require that process evaluation teams include experienced social scientists that use social science theory in the design of projects and analysis of findings.
- Encourage greater connection to the general evaluation community, which serves a wide variety of disciplines, and expect process and market evaluators to have training in evaluation.

References

- American Evaluation Association. 2003. *Guiding Principles for Evaluators*. Electronic copy downloaded 1/30/09: <http://www.eval.org/Publications/GuidingPrinciples.asp>. Fairhaven, Mass.: American Evaluation Association.
- Barata, S., and C. Anderson. 2007. "Making the Intangible Tangible: How to Evaluate A Social Marketing Campaign." In *Proceedings of the 2006 ACEEE Summer Study on Energy Efficiency in Buildings*, 862-69. Washington, D.C.: American Council for an Energy-Efficient Economy.
- Bensch, I., and S. Pigg. 2002. "Bird's Eye View of Energy Efficiency Market Research: Time to Move Beyond the Forest." In *Proceedings of the 2006 ACEEE Summer Study on Energy Efficiency in Buildings*, 10:23-34. Washington, D.C.: American Council for an Energy-Efficient Economy.
- 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*. Phoenix, Ariz.: Association of Energy Services Professionals.
- Dahler-Larsen, P. 1998. "Beyond Non-Utilization of Evaluations: An Institutional Perspective." *Knowledge, Technology and Policy* 11 (1&2, Spring/Summer): 64-90.
- Davidson, E. J. 2005. *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*. Thousand Oaks, Calif.: Sage Publications.
- Electric Power Research Institute (EPRI). 1998. *The High Efficiency Laundry Metering and Market Analysis (THELMA) Project*. Palo Alto, Calif.: Electric Power Research Institute.
- Forss, K., C. Rebien, and J. Carlsson. 2002. "Process Use of Evaluations." *Evaluation* 8 (1): 29-45.
- Gandhi, N., F. Kneipp, D. Moran, J. Peters, S. Samiullah, and A. West. 2007. "Product Selection – A Forgotten Vital Component of Program Design." In *Proceedings of the 2007 International Energy Program Evaluation Conference*, 1010-1021. Chicago, Ill.: International Energy Program Evaluation Conference.

- Grasso, P.G. 2003. "What Makes an Evaluation Useful? Reflections from Experience in Large Organizations." *American Journal of Evaluation* 24 (4): 503-514.
- Hall, N., M. Brown, P. Jacobs, and L. Megdal. (2005). "Overview of the Evaluations of California's 2002-2003 Efficiency Programs." In *Proceedings of the 2005 International Energy Program Evaluation Conference*, 41-52. Brooklyn, N.Y.: International Energy Program Evaluation Conference.
- Hirst, E., and J. Reed. (Eds.). 1991. *Evaluation of Utility DSM Programs*. Oak Ridge, Tenn.: Oak Ridge National Laboratory.
- Hodges, S.P., and M. Hernandez. 1999. "How Organizational Culture Influences Outcome Information Utilization." *Evaluation and Program Planning* 22: 183-197.
- Janda, K. B. 1994. "Bounded Decision Making and Analytical Biases In Demand Side Management." In *Proceedings of the 1994 ACEEE Summer Study on Energy Efficiency in Buildings*, 1:75-83. Washington, D.C.: American Council for an Energy-Efficient Economy.
- Joint Committee on Standards for Educational Evaluation. 1994. *The Program Evaluation Standards* (2nd ed.) Also available at <http://www.wmich.edu/evalctr/jc/>. Thousand Oaks, Calif.: Sage Publications.
- Kallock, W. B. 2005. "The Importance of Field Observations In Evaluations." In *Proceedings of the 2005 International Energy Program Evaluation Conference*, 389-392. Brooklyn, N.Y.: International Energy Program Evaluation Conference.
- Kane M., and W. Trochim. 2007. *Concept Mapping for Planning and Evaluation*. Thousand Oaks, Calif.: Sage Publications.
- Knowlton, L.W., and C.C. Phillips. 2009. *The Logic Model Guidebook: Better Strategies for Great Results*. Thousand Oaks, Calif.: Sage Publications.
- Kunkle, R., and L. Lutzenhiser. 1999. "Beyond Market Transformation: Some Perspectives On Energy Evaluation and Research and the Energy Efficiency Movement." In *Proceedings of the 1999 International Energy Program Evaluation Conference*. Denver, Colo.: International Energy Program Evaluation Conference.
- Lee, A. D., M.S. Khawaja, and M. Levy. 2007. "Using Simplified Pattern Matching to Define Program Theory and Assess Program Effectiveness – Application to a Green Buildings Program." In *Proceedings of the 2007 International Energy Program Evaluation Conference*, 622-633. Chicago, Ill.: International Energy Program Evaluation Conference.
- Lutzenhiser, L. 1993. "Social and Behavioral Aspects of Energy Use." *Annual Review of Energy and the Environment* 1993 18:247-89.
- Lutzenhiser, L. 2009. *Behavioral Assumptions Underlying California Residential Sector Energy Efficiency Programs White Paper*. Oakland, Calif.: California Institute for Energy and Environment.

- Mark, M. 2004. "The Mechanisms and Outcomes of Evaluation Influence." *Evaluation* 10(1): 35-57.
- Mast, B., and P. Ignelzi. 1996. "The Roles of Incentives and Information in DSM Programs." In *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings*, 10:145-153. Washington, D.C.: American Council for an Energy-Efficient Economy.
- Nicol, J., and C. Schepp. 2005. "Survey Results for the Practical Energy Management[©] Approach: Focus on Energy Program Response to Industrial Customer Feedback." In *Proceedings of the 2005 International Energy Program Evaluation Conference*, 497-508. Brooklyn, N.Y.: International Energy Program Evaluation Conference.
- O'Drain, M., and J.S. Peters. 1993. "The Process of Evaluation: Insights from Anthropology." In *Proceedings of the 1993 International Energy Program Evaluation Conference*. Chicago, Ill.: International Energy Program Evaluation Conference.
- Opinion Dynamics Corporation. 2008. *2006-2008 Statewide Marketing and Outreach Process Evaluation*. Study ID SCE0256.01. San Francisco, Calif.: California Public Utilities Commission.
- Owen, J.M. 1999. *Evaluation: Forms and Process*. Thousand Oaks, Calif.: Sage Publications.
- Peters, J. S., S.A. Baggett, P. Gonzales, P. DeCotis, and B. Bronfman. (2007). "How Organizations Implement Evaluation Results." In *Proceedings of the 2007 International Energy Program Evaluation Conference*, 35-47. Chicago, Ill.: International Energy Program Evaluation Conference.
- Peters, J.S., and M. McRae. 2009. *Process Evaluation Insights on Program Implementation*. Oakland, Calif.: California Institute for Energy and Environment.
- Peters, J. S., S. Nadel, and Z. Guijin. 2005. "China Motor Systems Energy Conservation Project Process Evaluation: A Cross-Cultural Process Evaluation of a Cross-Cultural Program." In *Proceedings of the 2005 International Energy Program Evaluation Conference*, 519-530. Brooklyn, N.Y.: International Energy Program Evaluation Conference.
- Peters, J.S., and B. Tannenbaum. 1997. "Market Simulation as a Method to Assess Market Potential for Solar Domestic Hot Water Program in Wisconsin." In *Proceedings of the 1997 International Energy Program Evaluation Conference*, Chicago, Ill.: International Energy Program Evaluation Conference.
- Rust, R.T., and R. L. Oliver. 1994. *Service Quality: New Direction in Theory and Practice*. Thousand Oaks, Calif.: Sage Publications.
- Shel Feldman Management Consulting, Research Into Action, Inc., and Xenergy. 2001. *The Residential Clothes Washer Initiative: A Case Study of the Contributions of the Collaborative Effort to Transform a Market*. Boston, Mass: Consortium for Energy Efficiency.
- Spinney, P., J.S. Peters, and P. O'Rourke. 1992. *DSM Process Evaluation: A Guidebook to Current Practice*. PR-100647. Palo Alto, Calif.: Electric Power Research Institute.

- Sullivan, M.J. 2009. *Behavioral Assumptions Underlying Energy Efficiency Programs for Business White Paper*. Oakland, Calif.: California Institute for Energy and Environment.
- TecMarket Works Framework Team. 2004. *The California Evaluation Framework*. San Francisco, Calif.: California Public Utilities Commission.
- TecMarket Works Team. 2006. *California Energy Efficiency Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*. San Francisco, Calif.: California Public Utilities Commission.
- Kane, M., and W. Trochim. 2007. *Concept Mapping for Planning and Evaluation*. Thousand Oaks, Calif.: Sage Publications.
- United Way. 1996. *Measuring Program Outcomes: A Practical Approach*. <http://www.unitedwaystore.com>. Alexandria, Va.: United Way of America.
- Valovirta, V. 2002. "Evaluation Utilization as Argumentation." *Evaluation* 8(1): 60-80.
- Vine, E. 2008. "Strategies and Policies for Improving Energy Efficiency Programs: Closing the Loop Between Evaluation and Implementation. In *Proceedings of the 2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 5:324-335. Washington, D.C.: American Council for an Energy-Efficient Economy.
- Weiss, C. 1998. "Have We Learned Anything About the Use of Evaluation?" *American Journal of Evaluation* 19(1): 21-34.
- Weiss, C.H., and M.J. Bucuavalas. 1980. "Truth Tests and Utility Tests: Decision Makers Frames of Reference for Social Science Research." *American Sociological Review* 45 (April): 302-313.
- Wilson, C., T. McDaniels, and R. Bennett. 2006. "What Developers Think About District Energy: A Mental Models Approach." In *Proceedings of the 2006 ACEEE Summer Study on Energy Efficiency in Buildings*, 11:147-158. Washington, D.C.: American Council for an Energy-Efficient Economy.
- Wirtshafter, R.M., and S. Samiullah. 2005. "Who Are the Real Non-Participants in California: A Two-Year Examination of California's Residential Energy Efficiency Portfolio?" In *Proceedings of the 2005 International Energy Program Evaluation Conference*, 1049-1060. Brooklyn, N.Y.: International Energy Program Evaluation Conference.