

Overview of the Evaluations of California's 2002-2003 Efficiency Programs

Nick Hall, Carmen Best, Johna Roth, TecMarket Works, Oregon, WI
Marian Brown, Southern California Edison, Rosemead, CA
Pete Jacobs, Architectural Engineering, Boulder, CO
Lori Megdal, Megdal & Associates, Acton, MA

ABSTRACT

In 2002-2003, California fielded eighty-five energy efficiency programs funded through the public goods charge (PGC). These programs included statewide, local utility, and non-utility implementation efforts. Each program was to be independently evaluated according to the California Public Utility Commission's (CPUC's) Energy Efficiency Policy Manual. This paper presents the evaluation approaches used to address the CPUC's evaluation objectives, the issues addressed by the various studies, the activities used to address those issues, and the level of program effects achieved by the programs as reported in the program evaluations. It also presents findings on the potential for consolidated research across the programs as well as some of the lessons learned by the CPUC in the CPUC's first comprehensive effort to manage the evaluations of a wide range of public goods charge programs.

Introduction

From 2002 to 2003 California fielded eighty-five energy efficiency programs funded through the public goods charge. Of these eighty-five programs, thirty-one were information or education programs, and the remainder represented a range of strategies to encourage adoption of energy-efficient technologies and practices. Programs were engaged in a variety of market sectors including the commercial, residential, local government, schools, industrial, and agricultural sectors. Several programs addressed multiple market sectors simultaneously through cross-cutting efforts. Evaluation methods also varied from one program to the next based on the CPUC-approved evaluation objectives for the specific program, the evaluation objectives of the program being evaluated, and the evaluation budget. This paper provides a high-level summary overview of the 2002 statewide IOU programs and the 2002-2003 local IOU and non-utility programs, the evaluation approaches used and some of the results from these studies. At the time of preparing this paper most (all but nine) of the 2002-2003 local IOU and non-utility or third-party program evaluations were completed and available for inclusion. Likewise, all of the statewide IOU 2002 program evaluations were available; however, none of the 2003 statewide IOU program evaluations were complete. As a result, the contents of this paper include the results from most of the local IOU and non-IOU 2002-2003 evaluations and the 2002 statewide IOU program evaluations. This consideration is important, as the energy savings information presented in this paper cover one year of statewide IOU programs and two years of non-IOU programs. Consequently, the energy savings should not be compared as being from equivalent periods of time or program durations.

Program Types, Target Markets, and Actors

There were primarily two types of programs implemented in California funded by the public goods charge: information and education programs (information) that provided information to the targeted market actors appropriate for each program and resource acquisition programs (resource) acquiring energy impacts through technology adoption approaches. These two types of programs can be sorted by the market sectors targeted by the programs. Figure 1 presents the market sectors targeted by both the information and resource programs.

The largest proportion of information programs attempted to reach multiple markets through cross-cutting efforts. Most of these programs were implemented by the utilities through both statewide and local delivery programs. Resource acquisition programs most often targeted residential and commercial markets, with the largest number of these programs being implemented by non-utility organizations. In dollar terms, most of both the utility and non-utility program funds went to resource programs (see), and most of total funds went to utility programs.

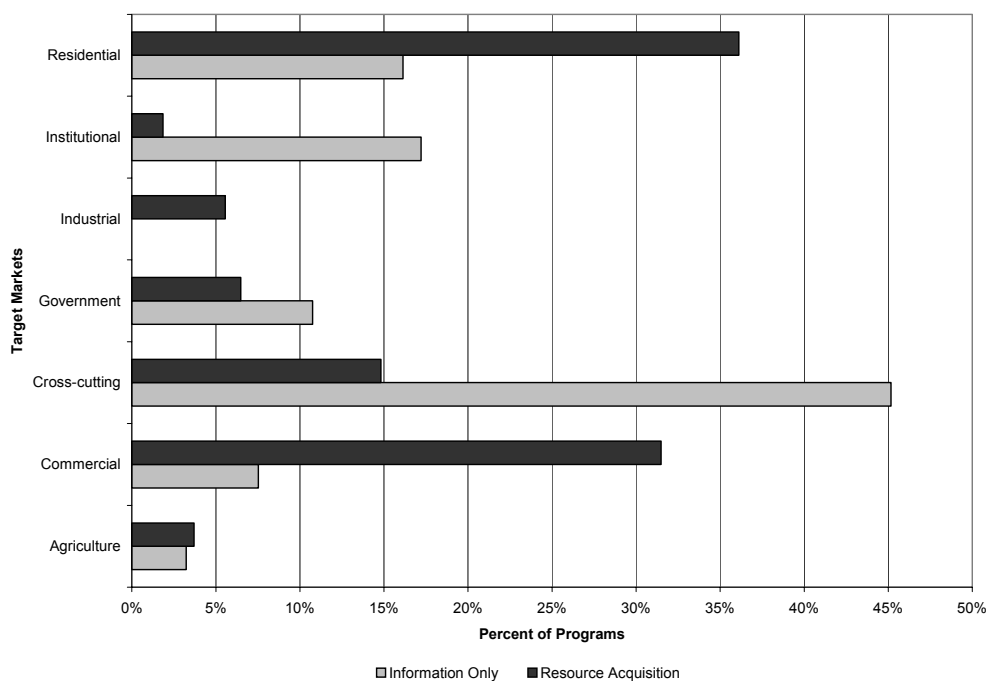


Figure 1. Distribution of Program Target Markets

Of the eighty-five programs implemented from the public goods charge, thirty-one programs were officially classified as information programs. Most of these programs and several of the resource acquisition programs also had educational components to them, but the identification as an “information” program meant that the CPUC did not require these programs to meet energy savings goals or conduct an impact evaluation. All other programs were obligated to set and meet established energy savings (resource acquisition) goals. A breakdown of California’s program portfolio by these criteria is presented in.

Table 1. California’s 2002-2003 Program Classifications

Program Classification	Number of Programs	Percentage of Programs	Total Budget*	Percentage of Budget
Information Only	31	36%	\$37,813,135	21%
Resource Acquisition	54	64%	\$142,249,413	79%
Total	85	100%	\$180,062,548	100%

* Includes 2002-2003 local IOU and non-IOU programs and 2002 statewide IOU programs.

Many programs focused on business owners and managers, but also made inroads with building managers, contractors and designers. The residential programs mostly targeted the single family, moderate income, and multifamily markets, targeting homeowners and hard-to-reach customers. (Low-income programs are budgeted and regulated separately.) A variety of other programs addressed a number of market segments including agriculture and industrial markets; others had a broad “cross-cutting” approach working in multiple market sectors and with multiple market actors simultaneously. Table 2 presents an overview of the market actors targeted by both information and resource programs. This table also identifies the number of programs targeting specific market actors and the percent of all programs that focused efforts on the identified actors.

Table 2. Targeted Market Actors*

Program	Actors Targeted	Number of Programs*	Percent of all Programs*
Resource Acquisition	Business manager / owner	28	52%
	Home owner	16	30%
	Contractor / builder	14	26%
	Renter	11	20%
	Multi-family: owner / manager	10	19%
	Manager: plant, facility, property, etc.	8	15%
	School official / administrator	5	9%
	Non-English speaking business managers / owners	4	7%
	Energy manager	4	7%
	Distributor	4	7%
	Waste water manager & operator	4	7%
Information Only	Business manager / owner	11	35%
	Building operators	9	29%
	Contractor / builder	8	26%
	Designer / architect / engineer	8	26%
	Home owner	7	23%
	School official / administrator	7	23%
	Manager: plant, facility, property, etc.	6	19%
	Energy manager	5	16%
	Students	4	13%
	Code official	3	10%
	Local government officials	3	10%
	Manufacturer	3	10%
	Renter	3	10%
	Chinese speaking residents	2	6%
	School teacher / instructor	2	6%
	Distributor	2	6%
	Multi-family: owner / manager	2	6%
	Growers using irrigation	2	6%
Waste water manager & operator	2	6%	

* Includes 2002-2003 local IOU and non-IOU programs and 2002 statewide IOU programs.

California's suite of 2002-2003 programs included four basic types of programs characterized by their implementation administrators. Three types of local programs were implemented through government organizations (GOV), investor-owned utilities (IOU), or third party implementers (TPI). Investor owned utilities were also responsible for several statewide (SW-IOU) programs.

Evaluation Approaches, Methods and Issues Addressed

This section of the paper provides an overview of the evaluation efforts.

The focus of the evaluation efforts varied across programs. This variation was primarily driven by the program's focus, the implementer's evaluation desires and the CPUC's approved evaluation objectives. However, the available supportive budget, in most cases, determined the rigor and comprehensiveness of the evaluation effort. Both information and resource acquisition program evaluations included process evaluation components. Logically, the resource acquisition program evaluations were more likely to include impact evaluations despite a few information program evaluations that also conducted energy impacts studies (See Figure 2).

As expected, the evaluation methods employed are closely related to the type of evaluation that was conducted indicates that while both resource acquisition and information program evaluations relied heavily on participant surveys, resource acquisition program evaluations also were more apt to use on-site inspections and audits to evaluate energy savings. However, while the CPUC Policy Manual required the use of IPMVP evaluation approaches (which require on-site measurements), many evaluations did not conduct on-site or site-specific verification, monitoring or metering activities. In fact, just over 50% of the impact evaluation efforts, in some way, complied with this requirement.

Information program evaluators took advantage of program staff interviews to inform their evaluation efforts. Understandably, qualitative evaluation methods were more prevalent in evaluating the information programs, while resource acquisition program evaluations were more apt to use quantitative methods to focus on energy savings evaluation.

Process Evaluation

The topics addressed in the process evaluations varied to a certain extent by the type of program offered, but many of these issues could be addressed within the process evaluation efforts for any program. In the 2002-2003 evaluation period, more leeway was given to the administrators to focus the investigative areas covered by the process evaluation than was provided in 2004-2005. In 2002-2003, program evaluations focused on program operations and implementation systems, marketing and outreach activities, overall satisfaction and reviews of the program tracking systems.

The percent of the process evaluations addressing each process topic is presented in Figure 4.

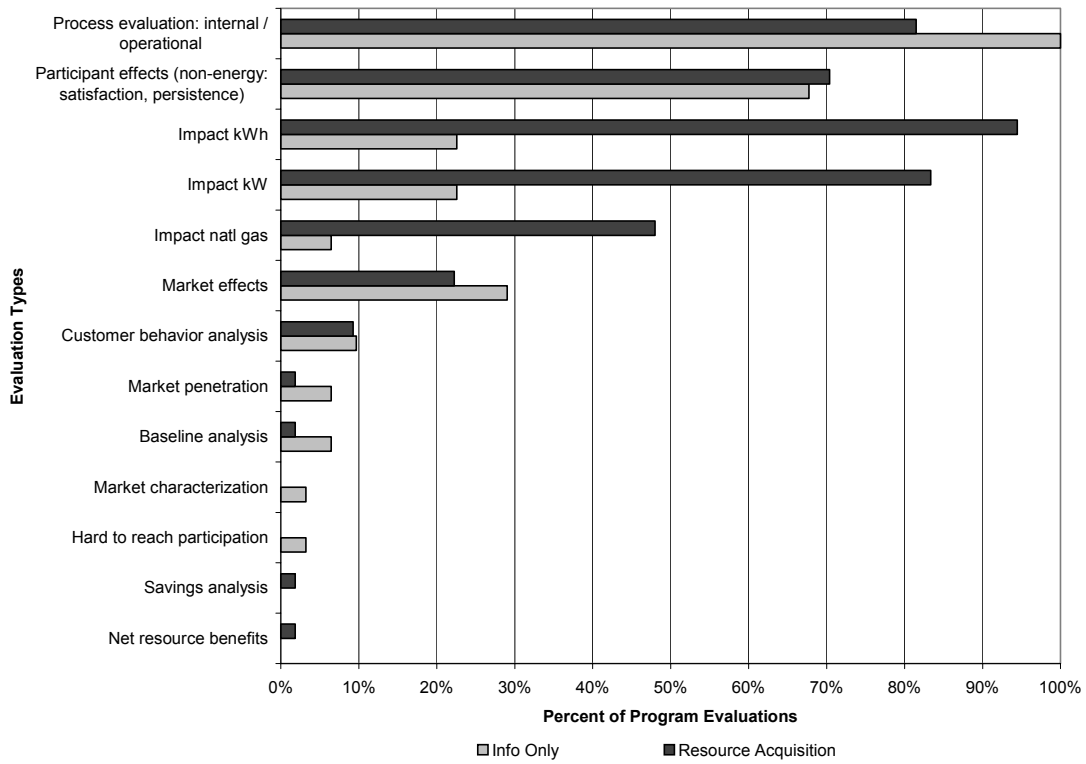


Figure 2. Types of Evaluations Conducted

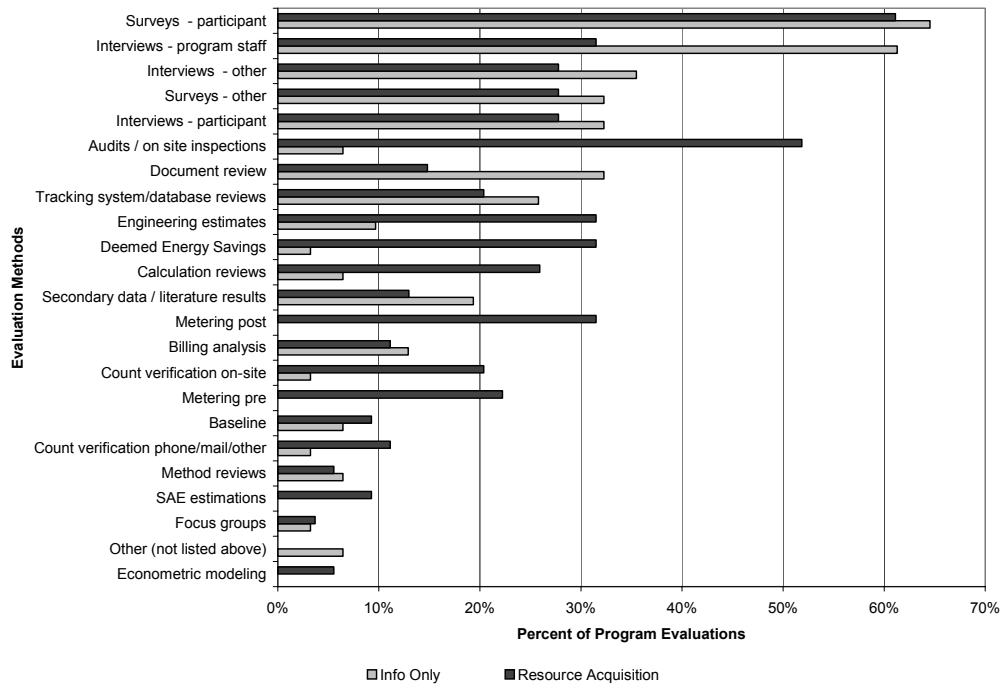


Figure 3. Evaluation Methods

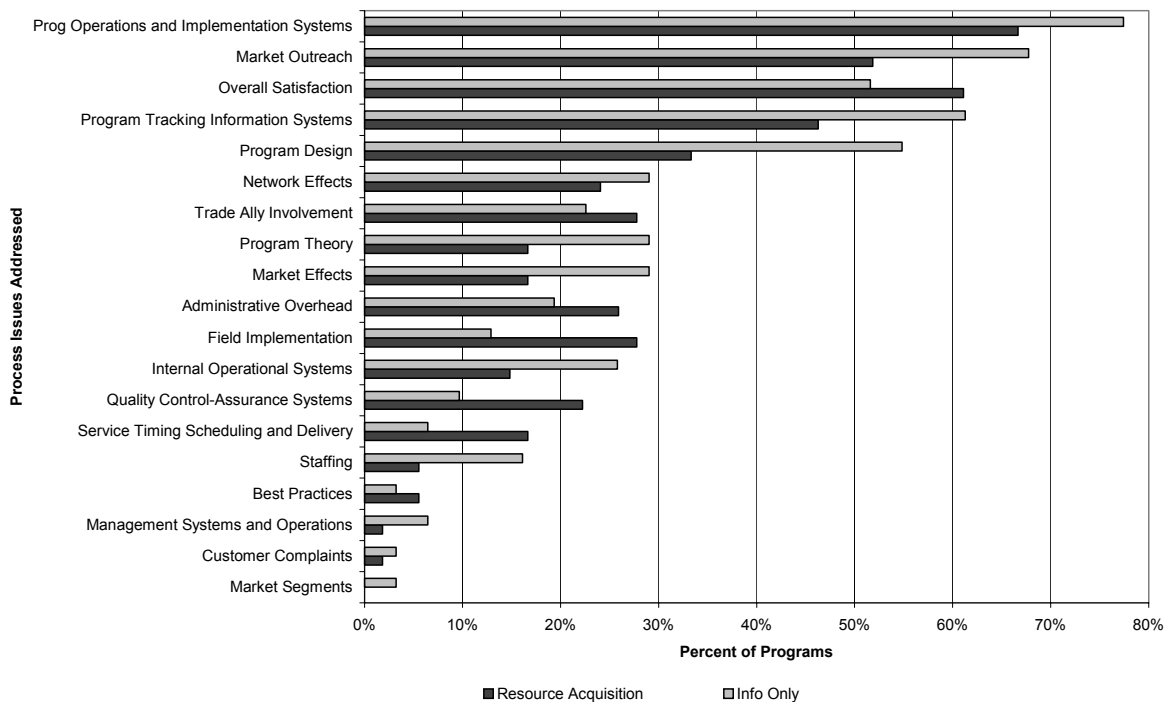


Figure 4. Evaluation Reports Addressing these Process Evaluation Issues

While the number of programs evaluating process issues is of interest and importance, the comprehensiveness of evaluating these issues is of even greater significance. For example, overall program satisfaction was a common evaluation criterion, but effective evaluation of this program component can take significant effort beyond a simple ranking of the general satisfaction question responses of high-level program components. A review of the satisfaction questions revealed that many evaluations focused on general satisfaction issues and did not go deeper to focus on issues at the level needed to make program improvements or to identify the drivers of low or high satisfaction levels. Likewise, no satisfaction research focused on the three primary relationships that influence satisfaction (customer-product relationships, product-provider relationship, and customer-provider relationship). As a result, only a few evaluations provided satisfaction information of value to program designers or were useful for understanding what program design or delivery changes could be implemented to improve customer satisfaction, one of the key purposes of the evaluation effort. Thorough analysis of participant satisfaction should address design and delivery issues from multiple perspectives and various degrees of detail to arrive at a measure of satisfaction that can be directly used to change program components or operational practices.

In reviewing program evaluations that included some form of participant satisfaction, we found that “overall” (i.e. overall, how satisfied are you with ...) satisfaction questions were the most common satisfaction measure. More indepth analysis of other components of satisfaction, like service, or product, or service delivery approaches, were less commonly assessed. In reviewing the evaluation reports, we identified three types of satisfaction that were evaluated: overall program satisfaction, satisfaction with the product they obtained, and satisfaction with the service provided by the program. In this assessment we found that 43 percent of the 56 evaluations that asked about satisfaction asked only one type of satisfaction question. Of this 43 percent, 88 percent only asked about overall program satisfaction, 8 percent only asked about satisfaction with the product they obtained and 4 percent only asked about service satisfaction. Thirty-one percent of the evaluations that asked about satisfaction

asked about two satisfaction components. Overall satisfaction again was the most common element to include, with 76 percent of these studies measuring overall satisfaction in addition to one other satisfaction measurement. Twenty-seven percent of the studies that asked satisfaction questions asked about all three of these satisfaction elements. Figure 5 presents these results graphically.

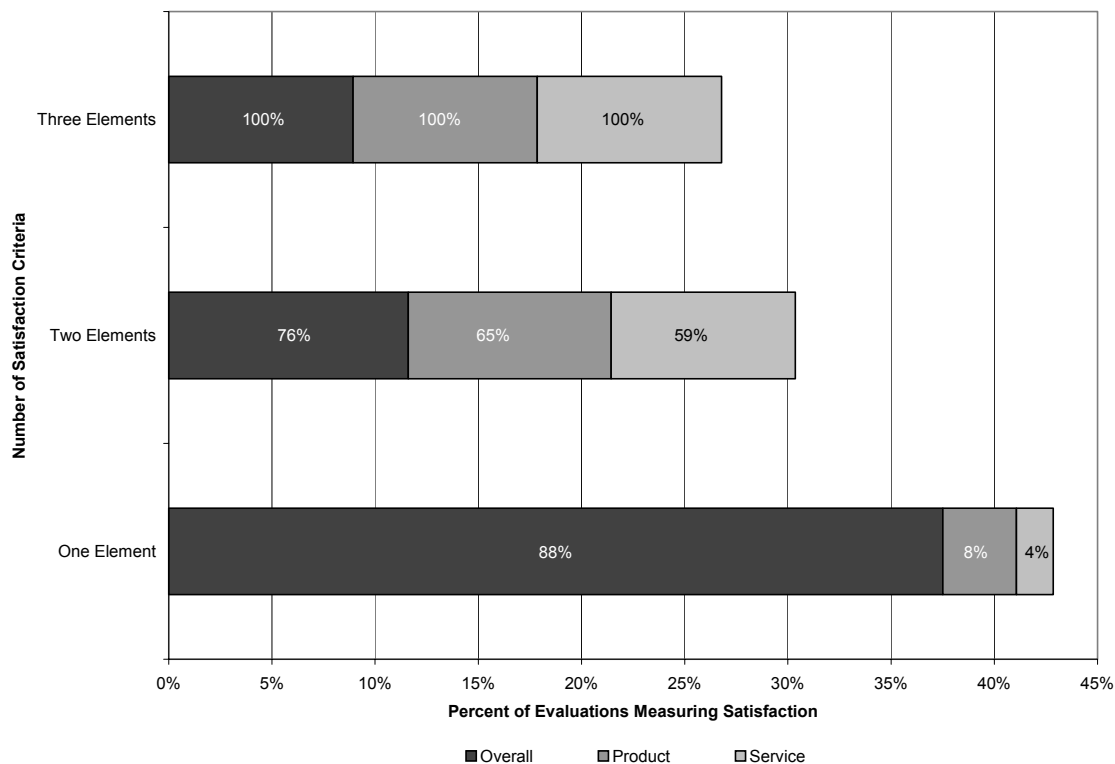


Figure 5. Comprehensiveness of Process Evaluation – Participant Satisfaction

One of the key goals of the process evaluation is to make recommendations to improve the programs being evaluated; this is especially true for resource acquisition programs, for which California depends on the program’s impacts to meet the State’s energy needs. Figure 6 and

Figure 7 present the process issues examined in the evaluations and indicate if the evaluation made a recommendation on that issue that was designed to improve the program. Across all of the process evaluation issues examined, if the evaluator looked into a particular issue, recommendations for improvements were made over half the time.

However, the issue with the way satisfaction was measured, as described earlier, demonstrates the weakness of the evaluation approaches. For information programs that examined satisfaction, only about 25% of the studies provided recommendations for improving satisfaction. For resource acquisition programs, a little over 10% of the evaluations made recommendations to improve satisfaction. An optimistic assessment of this data is that these programs have reached their satisfaction zenith and only 10% of the acquisition programs are capable of improving satisfaction. A pessimistic assessment is that the satisfaction measurement approaches used by the evaluations were inadequate to provide the detailed information needed to assess satisfaction and develop recommendations. A review of the questions asked indicates that the latter may be closer to reality.

There is clearly a need for evaluation professionals to understand how to design and ask satisfaction questions that can lead to recommendations to improve program performance. On the other hand, evaluators are accustomed to investigating the operations and operational systems of a program and developing recommendations to improve program operations. Just over 70% of the information program evaluations and over 80% of the resource program evaluations that examined operational components made recommendations to improve program performance.

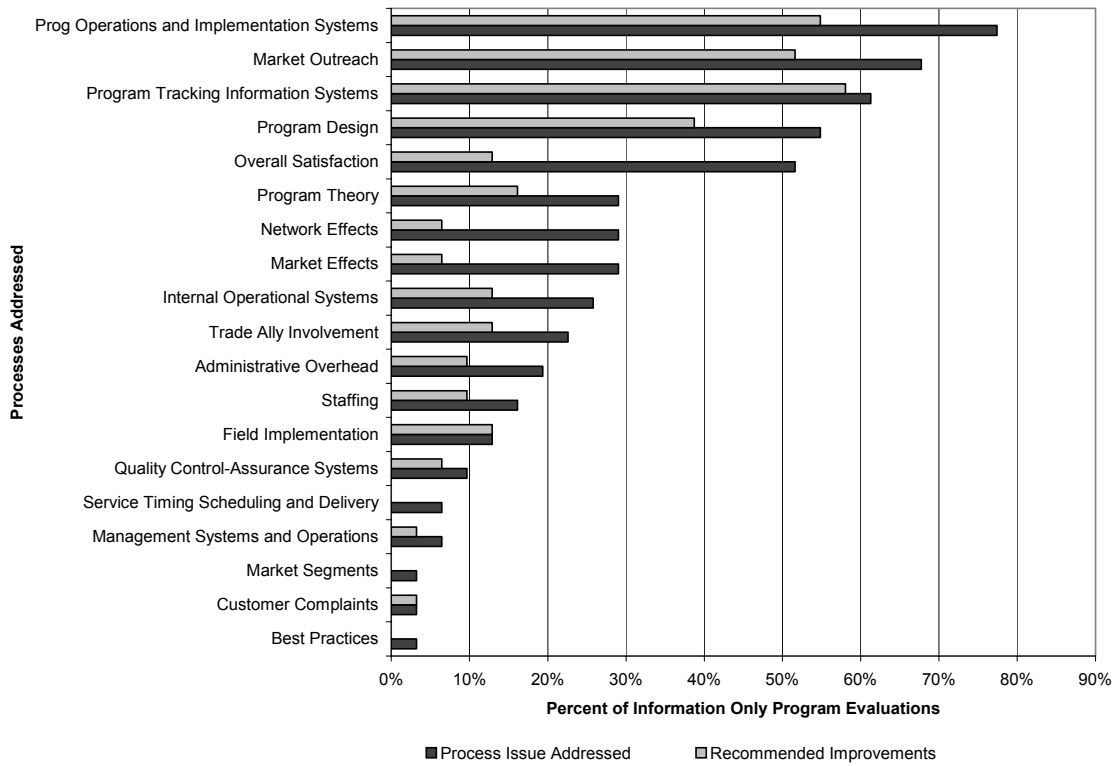


Figure 6. Information Program Evaluations Making Recommendations to Improve

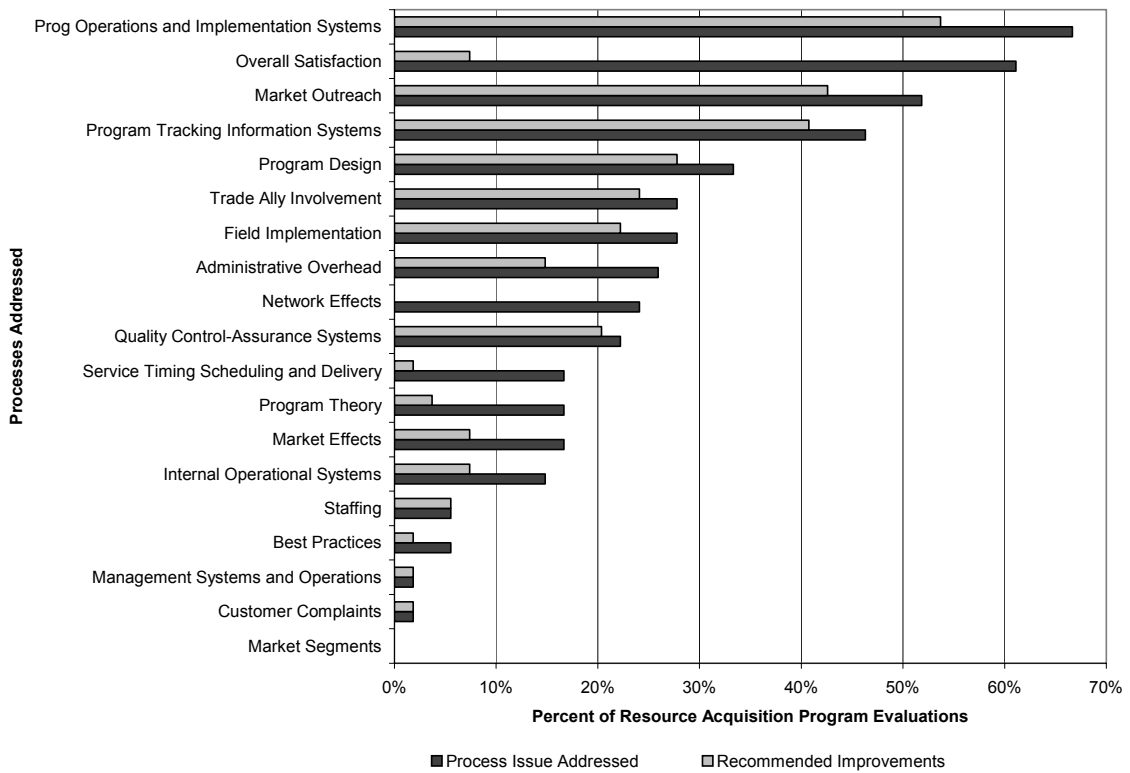


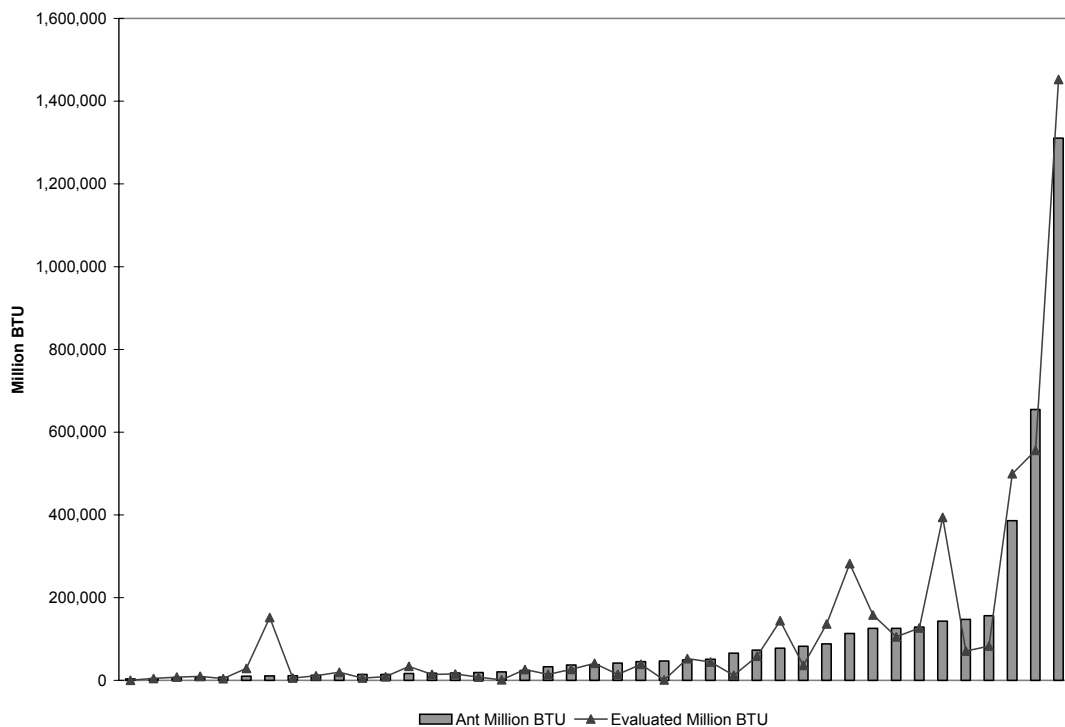
Figure 7. Resource Program Evaluations Making Recommendations to Improve Measuring Energy Impacts

The following analysis is based on a review of the 2002 statewide IOU program evaluations and the 2002-2003 local IOU and non-IOU program evaluations.

Based on acquisition program implementation plan savings targets from the suite of California's 2002 statewide IOU and 2002-2003 local IOU and TPI programs, the portfolio was anticipated to save 1,150,019,889 kWh and 24,966,162 therms. Of the originally anticipated savings, 788,849,882 kWh and 15,650,943 therms were evaluated to some degree in reports completed as of this paper. However, at this time nine studies have not been reviewed for the 2002-2003 TPI programs and none of the 2003 statewide IOU program evaluations are completed. Therefore these results are not included in these estimated savings. The program evaluations reviewed for this paper cover about 66 percent of the total portfolio's projected savings.

The savings reflected in the evaluation reports received to date indicate that the programs saved 808,924,508 kWh and 19,497,320 therms. These studies indicate that the programs achieved their original energy saving targets, suggesting that the suite of programs somewhat under-estimated their savings impacts. These underestimations came from the investor-owned utility programs, both local and statewide. That is, the IOUs seem to outperform their savings targets. In addition, because the evaluations may not have reported all of the savings for the evaluated programs because of the need to conduct the evaluations during the implementation period, the actual achieved savings may be greater than those reported in this paper.

The comparison between the evaluated energy savings and the energy savings targets (anticipated energy savings) approved by the CPUC is presented in Figure 8. The programs for which the evaluated savings line is greater than the anticipated savings (Ant) are all IOU- managed programs.



Demand reduction is an important component of energy efficiency programs, but it was not universally reported in the evaluation reports. Thirty resource acquisition program evaluations reported demand savings (71 percent of programs with evaluations), as did three information programs. Lifetime energy savings are much less commonly evaluated and annual load impacts are not reported in any of the studies. Only 12 out of 42 resource acquisition program evaluations reported lifetime energy savings (29 percent). Clearly there is a need for minimum reporting standards from the evaluation efforts, if these evaluations are to effectively inform public policy on energy resource planning.

In Figure 9 we present the frequency with which the impact evaluations employed standard evaluation techniques to adjust the estimated savings to improve the accuracy of the estimate. Note that 46 percent of the anticipated energy savings were evaluated using verification of installation approaches, even though IPMVP on-site field protocols were required of all evaluations. Non-verification field information of various types was also used to evaluate and adjust 37 percent of the anticipated savings. About 35 percent of the evaluated savings were adjusted for free-ridership, 6 percent were adjusted for participant spillover, and 4 percent were adjusted for non-participant spillover. Two percent of the savings were adjusted for education effects and market effects. It is clear from these findings that the program impacts presented in the evaluation reports are not comparable across the studies conducted. The savings calculations were not consistently conducted across the portfolio, which clearly indicates a need for detailed evaluation protocols to guide the evaluation and impact accounting efforts.

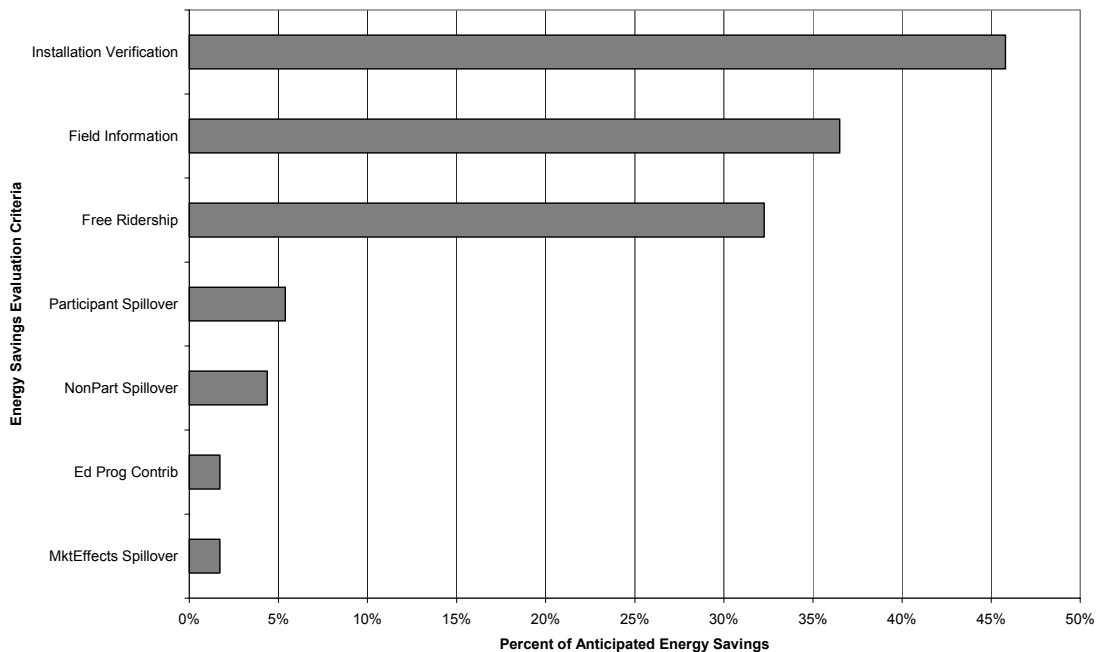


Figure 9. Energy Impact Evaluation Methods

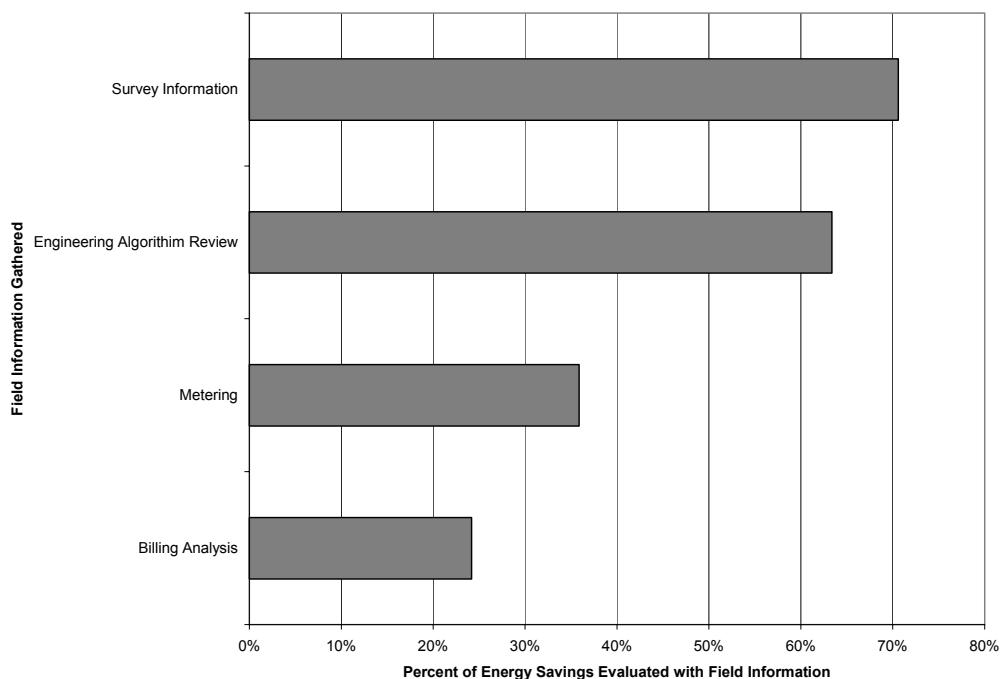


Figure 10. Field Methods Used to Adjust Energy Savings

Just as different calculation approaches were used (see above), the evaluations also employed a number of different types of tools in the estimation process. Figure 10 presents the commonly used tools to assist in the estimation process. The most commonly used tool was the use of surveys. In the evaluations covering 71 percent of the energy savings, evaluators used survey data to help estimate impacts. About 64 percent of the savings were evaluated using engineering algorithm reviews of program-provided estimated savings, about 36 percent of the savings were evaluated using evaluation metering data, and about 24 percent were evaluated using some form of utility metering analysis (commonly referred to as billing analysis).

Lessons Learned

In conducting the duties of the 2002-2003 California Master Evaluation Contract, the TecMarket Works team learned a number of valuable lessons for future consideration. A sample of these lessons learned includes:

1. Help the implementers and evaluation contractors, particularly the TPI contractors, understand the evaluation requirements of the Energy Efficiency Policy Manual, especially what is required to comply with IPMVP field procedures (International Program Measurement and Verification Protocols) so that they understand that all IPMVP options require some form of on-site field measurements.
2. Establish clear definitions of the evaluation metrics that are required from the evaluation process, including definitions for energy savings (kWh), demand impacts (kW), lifecycle or lifetime savings and evaluation-confirmed, net-adjusted savings.
3. Help the budgeting process understand the complexity of the evaluation process and what resources are required to meet Policy Manual and IPMVP requirements.

4. Establish clear evaluation protocols that evaluations must follow, especially protocols for reporting and estimating net ex-post energy impacts and for conducting process evaluations.
5. Provide clear guidance on how to conduct satisfaction assessments so that the assessments directly lead to program improvement recommendations and can be comparable across programs.
6. Help implementers and evaluators understand what the term “arms-length evaluation effort” means.
7. Provide help to implementers so that they know what skills are required of the staff that is conducting the evaluation studies.
8. Establish impact estimation protocols and sampling requirements so that savings estimates can be both reliable and representative.
9. Require that all evaluations present the level of precision and the error bands around their estimates.
10. Provide adequate time for the evaluation contracting and planning process.
11. Require or financially encourage evaluators to submit reports on time.

Consolidation Potential for Future Evaluation Efforts

In conducting the reviews of the 2002-2003 evaluation efforts, there appeared to be substantial potential for evaluation cost savings through consolidated evaluation efforts that would improve evaluation accuracy, reliability and comparability. These potentials are presented in a consolidation report¹ but in general they include consolidation of the evaluation effort by:

1. Program implementer, so that all programs implemented by a single implementer or a group of implementers are conducted together (IOU programs, IOU-managed programs, third party programs, non-profit providers, local governments, etc.);
2. Type of program, so that similar programs are evaluated as a group (audit programs, information and education programs, partnership programs, technology incentive programs, etc.);
3. Target market sector, so that programs focusing on the same market can be evaluated together (hard to reach, residential, small commercial, large commercial, industrial, agricultural, etc.);
4. Type of technologies offered (compressed air, HVAC tune-ups or upgrades, etc.).

Each of these consolidation approaches has a different set of benefits and challenges associated with the approach. These are discussed in more detail in the referenced report. However, one of the key issues that cut across all consolidation approaches is the ability of the evaluation industry to implement large-scale evaluation efforts. Most evaluation firms are small, employing only a few key evaluation professionals. When large evaluations are conducted, these firms must typically group into larger evaluation teams. These approaches require exceptionally skilled and experienced evaluation directors and managers to allow the evaluation efforts to succeed and to achieve the benefits of the consolidation approach.

¹ TecMarket Works 2002. “Recommendations for Evaluation Consolidation of California’s Energy Efficiency Programs.” October 30, 2002.