Are We There Yet: Building Behavior Programs to Serve a Purposeful Role in DSM Portfolios

Dr. Shahana Samiullah, Southern California Edison, Rosemead, California
Dr. Miriam Fischlein, Southern California Edison, Rosemead, California
Ms. Patrice Ignelzi, Applied Energy Group, Walnut Creek, California

Abstract

A new breed of efficiency initiatives, termed behavior programs, is very much on the minds of program planners and policy makers. Many such programs are in operation in North America. The programs and the policy mandates for inclusion of these programs in DSM portfolios need to first address a myriad of questions, including their purpose, role, cost-effectiveness, and potential savings in DSM resource portfolios.

In this paper we articulate some of the hard questions California is asking of such programs. Can behavioral programs be relied on for system planning or addressing grid constraints? Do they have a role as resource acquisition programs or should they be leveraged and serve as a complement to traditional programs? The experience from California includes sharing information about programs designed to fulfill regulatory mandates. In particular, the paper presents experience from one Southern California utility on program design considerations for feedback programs during planning and implementation phases, and on how evaluation was built into those designs to address the hard questions. While there is a variety of feedback programs that fall under an expanded category of behavior programs[1], this paper discusses role of these programs with a focus on the recent widespread adoption of home energy reports programs and similar usage feedback programs.

Introduction

What's novel about behavior programs in a DSM portfolio? The question is as interesting as its answer and many policy makers around the world are becoming well acquainted with this seemingly new, innovative breed of programs, termed behavior programs, in the traditional DSM portfolios. Basic research[2] and practitioner implementation in this area have proliferated in recent years, with close to 300 utility led programs being offered between 2008 and 2013.[3] However, there are many questions that still need to be raised and addressed as portfolios pursue inclusion of these programs, including their purpose, role, cost-effectiveness, and future potential in DSM resource portfolios. In this paper, we discuss the experience with behavior programs as program planners and policy makers attempt to address the critical questions about these programs and the path toward including them in the portfolio in California. We pose several critical questions for program planners, program evaluators, and policy

[1] Ignelzi, P.; J. Peters; K. Randazzo; A. Dougherty; L. Dethman; L. Lutzenhiser (2013)."Paving the Way for a Richer Mix of Residential Behavior Programs." Prepared for the California Investor-Owned Utilities: Pacific Gas & Electric, Southern California Edison, San Diego Gas & Electric, and Southern California Gas. San Francisco, CA. CALMAC Study ID: SCE0334.01. May 2013.
http://www.calmac.org/publications/Residential_Behavior_White_Paper_5-31-13_FINAL.pdf

[2] Delmas, M. A., Fischlein, M., & Asensio, O. I., (October 2013). Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy, Volume 61*, Pages 729-739.

[3] Mazur-Stommen, Susan & Farley, Kate (2013): ACEEE Field Guide to Utility-Run Behavior Programs. http://www.aceee.org/research-report/b132

makers using a concrete example of one type of program—comparative energy usage feedback programs—to illustrate their importance.

We decided to draw on recent evaluation experience with this type of program to both pose and help address key questions during the various phases of implementation of such programs in their varied roles as meeting goals, mandates, and resource acquisition strategy. We hope to help the various stakeholders interested in energy-related behavior programs and behavior-change strategies in DSM from direct experience from Southern California Edison (SCE) as it addressed the inclusion of such programs in its DSM portfolio. To the policy makers and program implementers, the paper provides examples from California on the possibilities for inclusion of these programs in their varying roles in a portfolio. For third-party evaluators and implementers, the paper identifies the role they can play in the development of programs, to achieve the required robustness of program design and ensure these programs' purposeful role in a DSM portfolio.

The paper lays out the important decisions that need to be made while considering the design of behavior programs and their inclusion in a DSM portfolio. In doing so, it also provides the practical experience of SCE when it made those decisions. And finally, based on the experience in California, we make some conclusions and recommendations for inclusion of behavior programs in a DSM portfolio.

Considerations in Behavior Program Design and Inclusion in DSM Portfolio

As plans are made to include behavior programs in the DSM portfolio in California, program planners and policy makers are working on what these programs should look like and how they will function there. At the same time, various programs are under design, in operation, and under evaluation for the impacts they have on energy use. Based on this experience with both of these paths, we see that there are four broad considerations for future design and inclusion of behavior programs in the DSM portfolio.

Designing and Operating the Behavior Program

Before considering whether to include in the portfolio, program planners must decide what type of program to consider and when it is ready to move from pilot stage to full-scale deployment. Here we discuss factors in those design and deployment decisions

- Use off-the-shelf or design your own program?
- Program to fulfill regulatory mandate or discretionary offering?
- What different perspectives do program managers, implementers, and evaluators have?
- Part of or in parallel with other DSM programs?

Use off-the-shelf or design your own program?

In most of the US states and for the most part in California as well, program administrators have implemented comparative energy usage feedback programs using off-the-shelf offerings.[4] As a result, a whole assortment of vendor offerings has emerged to fulfill the varied needs of behavior program inclusion in DSM portfolios. Based on authors' observations, there are both pros and cons of choosing between the two options.

[4] For example, third party implementer Opower alone works with 60 utilities to implement an off-the-shelf solution, while only seven utility-led behavior programs were counted by ACEEE before 2013. Mazur-Stommen, Susan & Farley, Kate (2013). ACEEE Field Guide to Utility-Run Behavior Programs. http://www.aceee.org/research-report/b132

Table 1. SCE Experience: Pros and Cons of Developing Custom/In-house Behavior Programs

Pros	**Cons**
• Better control of data analytics and insights as utilities are able to leverage their customer data better to develop business goals (e.g., sequencing of customers from one offering to another) • Greater understanding of embedded costs of such programs as end-to-end process is clearly known to determine the least cost program savings strategies • Ability to implement a customized approach to serving customers with co-marketing of such programs as an integrated DSM offering to achieve multiple objectives.	• A steep learning curve to integrating various internal systems and processes, and the time inefficiencies involved, • Aligning competing goals such as meeting yearly savings targets versus long-term investment for building the necessary resources and skills to implement these programs

Southern California Edison took a parallel path of adopting both off-the-shelf products while also investing in enhancing its existing program offering. Under California state law (D.12-05-015) investor-owned utilities have to reach a 5% behavioral program target for residential households by 2014. Behavioral programs countable towards this goal have to provide "comparative energy usage information," defined by SB 488 as programs "…pursuant to which an electrical corporation or gas corporation discloses information to residential subscribers relative to the amount of energy used by the metered residence compared to similar residences in the subscriber's geographical area." SCE is meeting part of its 5% mandate with OPower Home Energy Report. It provides SCE with relative simplicity Results of the evaluation we recently completed show definitive savings by its participants during the program year. But the utility had no say about the messaging, targeted savings tips provided, or access to the data analytics. To counter this and take advantage of the benefits noted above, SCE is also implementing a custom program designed and deployed in-house. The Custom Energy Reports pilot is a program with a more nuanced goal. In addition to showing participants how their monthly energy use compares with their neighbors, the program aims to educate them about how their usage relates to the tier pricing. Tier pricing means that kWh used in the first block, say 200 kWh, cost less per kWh than those in the next block. In SCE's tier pricing structure, the per-unit price in upper tiers is more than twice that in the lowest tier. In these reports, customers see graphs that show tier pricing affects their bill. For example, Figure 1 shows a customer whose bi-monthly usage was about the same in both the lower and higher tiers. The portion of the customers cost, however, was considerably higher for the higher tier usage.

Figure 1. Additional Customized Information Provided in Custom Energy Report

This information on the effects of tiered pricing, along with the participants' own 12-month history of usage in lower only and lower plus higher tiers, is included in addition to the comparative energy usage information of other similar customers. In those comparisons as well, participants are shown the effects of tiered pricing because they are compared with customers who have characteristics just like their own—similar home size, home vintage, weather, pool/no pool, low-income rate qualification, and bill consumption dates—except that the comparison customers have rarely reached the highest tier of use in the past year.

Program to fulfill regulatory mandate or discretionary offering?

In California, all investor-owned utilities responded to a mandate from California state legislature to offer a pilot behavior program.[5] Since then the four California IOUs have offered comparative energy usage disclosure behavior program on a limited scale to also fulfill another mandate from the Public Utility Commission: reach at least 5% of residential customers with behavior programs. This mandate has now been in place for two program cycles and expected to remain in place at least through 2015. Simultaneously, various research on behavior programs[6,7] identified the benefits of these programs, but also pointed out that more than one strategy for behavioral interventions might be worth experimenting with.[8,9] This prompted the IOUs to include similar programs as an innovative component in DSM portfolios. At the time these mandates were issued and then latter went into effect, only an offering from one third party vendor clearly satisfied the Senate bill requirements. Since that time,

[5] California State Senate Bill 488 specified requirements for electric and gas corporations to provide comparative energy usage disclosure information. The requirements would become inoperative on July 1, 2015, and would repeal on January 1, 2016.
[6] EPRI (2009). Residential Electricity Use Feedback: A Research Synthesis and Economic Framework. Report. Electric Power Research Institute: Palo Alto, CA.
[7] Delmas et al., 2013.
[8] Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency*, 1(1), 79-104.
[9] Ignelzi et al., 2013.

additional vendors have developed program designs that meet the requirements. And SCE, for one, has developed its own program. This means that the utility can meet its 5% requirement using a combination of programs that employ slightly different designs. While the mandated comparative feedback provision drove the early program pilots in California, the recent response to further research indicates that these strategies are "here to stay" in a successful DSM portfolio. In California and other states, notably New York and Wisconsin, there is considerable momentum to develop and test programs using alternative strategies, in the form of pilot programs implemented on a small scale. These are discretionary initiatives. But they can, and in some cases are intended to, influence regulatory policy about the types of programs eligible for inclusion in the portfolio. The timing, appropriate placement, who and what to target, are the important considerations for a purposeful role of these programs in a DSM portfolio.

Considerations utilities are beginning to address in these pilots, well before full-scale operation and inclusion in the portfolio, include: What intervention or combination of interventions are most effective and most effective on selected portions of the market? What's the most effective loading order of the strategies? Is the program most effective in inspiring new energy-saving behaviors or reinforcing or boosting ones encouraged by programs already in the portfolio? And which strategy(ies) seem to result in the most long-lasting or persistent savings?

What different perspectives do program managers, implementers, and evaluators have?

While IOU program managers initially considered these programs as mandates to be fulfilled and implementers responded with differentiated program offerings, the evaluators' circle started to ask the hard questions on 1) double-counting of savings when actions resulting from a behavior program treatment leads to participation and savings for other DSM programs in the portfolio, 2) persistence of savings beyond the treatment period and/or effective useful life of a treatment, and 3) causal linkages of postulated behavior-change theory behind the action-taking by customers for insights into treatment changes. The dialogue continues among these three perspectives as each tries to understand the issues posed. The dialogue has moved beyond proving the viability of behavior programs to understanding how these programs actually work for on-going program innovations. This is well demonstrated by the fact that for one California IOU, comparative energy usage program is the top energy savings measure in its current residential portfolio – a trend that is being identified to be rapidly spreading across many other US states residential energy efficiency portfolios.

What we are starting to see now is something that took years to evolve with traditional DSM programs—program planners and managers are consulting evaluators on designs that will best allow them to measure the savings of the behavior program. In addition, we are seeing up-front inclusion of metrics and evaluation methods for measuring program impacts by program planners and implementers. By establishing this flow of ideas, evaluation considerations matter to planners and implementers, and program designs and deployment push evaluators to develop ways of assessing program impacts. The world of social science behavior theories and that of evaluation methods and standards meet earlier, potentially cutting short the trial-and-error approach used for many years by DSM program administrators.

Part of or in parallel with other DSM programs?

Given the increased and prominent savings contribution of these residential behavior programs, many fear these programs are being used to capture only short term savings goals. In fact California Public Utilities Commission remains carefully balanced in its treatment of behavior programs and in providing further guidance on these programs. In a recently published industry conference paper, the Commission staff raised the concern about how these programs need to satisfy the long-term savings

aim as envisioned by the Commission to come from "deeper" retrofit projects and a "whole building" approach.[10]

Whether "behavior programs" should stand on their own or be strategies used to enhance traditional DSM programs for deeper and longer term savings is a question that may have more importance for regulatory categorization than in implementation or presentation to customers. As an administrative consideration, it's fair enough. In California, as in many states, energy efficiency (EE) programs and demand response (DR) programs have been segregated, though they co-exist in the portfolios of most utilities. In many jurisdictions the distinction hinges on EE's focus on kWh or energy savings and DR's focus by definition is on kW or demand savings. In the case of behavior programs, the distinction is less clear. Consider a traditional EE program—appliance rebates. The program encourages the purchase and installation of more energy efficient equipment. Through the years, they have demonstrated that these actions indeed result in energy savings. But it is not clear that the savings are all they can be. In Table 2, we illustrate that incorporating behavior program intervention strategies for water heaters to influence behaviors associated with hot water use could yield additional savings. These same behavior strategies could work as effectively as stand-alone behavior programs or as part of existing programs. In fact, these additional usage-related behaviors might be at work in the comparative energy and feedback programs in operation now. More on that below.

Table 2. Example of Combined EE/Behavior Program Designed to Change Appliance Purchase and Use Behaviors[11]

	Interventions	**Behavior Change**
Traditional Program Strategy	Financial incentives to consumers reduce price differential between standard and high-efficiency models, increasing demand for more efficient models	Greater availability and variety of more efficient water heater models available to consumers Replace storage water heater with more efficient unit when needed Encourage early replacement of water heater
Behavior Program Enhancements	In-person interactions (e.g., persuasion by a trusted contractor, word-of-mouth, behavior modeling through demonstration projects) Frame information about energy the way customers do and encourage follow-through	Get regular AC maintenance through a contract Buy a more appropriately sized unit Install low-flow showerheads and faucet aerators, too Turn down water heater temperature Drain sediment from storage tank for optimal performance Take shorter showers

[10] Samiullah, Shahana, Peter Franzese and Cathleen Fogel, Patrice Ignelzi (2014)."Integrating Behavior-Based Programs into the Portfolio." paper presented at 2014 Association of Energy Services Professional (AESP) Conference.
[11] Adapted from Ignelzi et al., 2013.

Assessing Reliability of Behavior Program Savings for Local Reliability Needs

It's difficult to consider including any program in the DSM resource acquisition portfolio without a high degree of confidence that the savings are reliable. In California the DSM portfolio is counted as a reliable resource that can offset need for capacity expansion. At SCE, the utility's distribution energy plan is including DSM programs among its distributed energy resources, which are being evaluated for locational benefits. The DSM resource is evaluated based on reductions in local generation capacity needs, avoided investments in distribution infrastructure and reliability benefits, as well as any other savings to the grid or costs to ratepayers. For behavior programs to be relied on for system planning or addressing grid constraints, they must be able to demonstrate savings that meet the established standard of reliability—both for magnitude and persistence of the savings. This is where measurement and verification become paramount to both 1) demonstrate savings {ex-post evaluations; Opower demonstration of savings during deployment year}, and 2) assess savings persistence. While the first evaluation aspect has been thoroughly measured over multiple program samples, there is not much data available on the degree to which those well-measured overall savings can persist to continue with or without the program and what hours of the day those savings are happening. The latter especially, is a key data need for using these programs for grid load management techniques. The quick ramp up of savings observed after initial treatment may indicate the instantaneous reaction of customers to the stimuli, but poses the question of what hours of system peak to rely on this stimuli and how to make this behavior change more habitual to address persistence for greater valuation of these resources.

Identifying What's Driving Behavior Program Savings

As noted above, the comparative feedback programs operated by Opower, and other similar programs, consistently demonstrate savings of 1-2% during the year they are deployed. We are only starting to understand how and when these savings are happening, and how these are persisting.[12,13] Part of the reason is that little information yet exists about the actions participants in these programs take that generate these savings. We need to pull back the curtain to identify the actions and program features that influence them.

The savings can be driven by any or all of the following factors, to name a few:

- Due to program targeting? Different behavior-change intervention strategies (such as comparative usage feedback) suggested by theories of behavior exist in part because it has been observed that different people respond to opportunities and choices differently. Interventions that resonate with some people might not move others to action at all. Part of what pilot program initiatives can do is test the effects of deploying the same program to different target markets. In SCE's case, Opower is implementing the same program to a different set of customers within the territory this year than last. Might the savings from these two groups be different and/or the actions behind them similar or different? Follow-up surveys in which additional customer characteristics are noted, along with actions they report having taken, could be conducted to provide insight into the variability among and relative suitability of the program in different target markets. In SCE's Custom Energy Reports program, the initial target market is especially high-usage customers, expected to be most responsive to the tiered pricing information in the reports. In subsequent years, the

[12] Steg, L. (2008). Promoting household energy conservation. Energy Policy, 36(12), 4449-4453.
[13] Asensio, A. & Delmas, M. (2014). The dynamics of information framing: The case of energy conservation behavior. UCLA working paper.

program could be tested on moderately high energy users to see if the reports elicit similar/different response and why.

- Due to program messaging? There is a whole body of social science behavior theory literature that says the way we present an idea affects how it is received and what we do in response.[14] In the off-the shelf Opower Home Energy Reports program, SCE has no control over and is not privy to the different messaging used and therefore cannot assess the effects. By contrast, in the Custom Energy Reports program that SCE has designed, several types of messaging about the effects of high energy use are being tested. Specifically, the experimental design will allow testing about the efficacy of loss prevention versus gain messaging. An example of alternative messaging being tested is
    - Loss prevention framing: *"You may be losing up to $150 per year by not reducing your energy use."*
    - Gain potential framing: *"You could save up to $150 per year by reducing your energy use."*

  As part of the program design, participants were randomly assigned to either the loss framing group or the gain framing group. A comparative analysis of savings in the two groups will be performed in the evaluation.
- Frequency of treatment? Does the size of savings attributable to the program depend on how many times a participant receives a report (aka the treatment)? Comparative feedback programs are commonly designed to deliver reports every couple of months. And while early evaluations suggest that reports delivered at 2-3 month intervals do yield significant savings that seem to persist throughout the program year, there is a dearth of publicly available research that has tested the relative effectiveness of different report frequencies. Is receiving 6 reports significantly more effective than 1 report? Is 6 more effective than 5? Does the number of reports affect the types of actions taken, the persistence of their effects, or perhaps both? Similar to testing alternative messaging, participants could be assigned to different frequency groups and a comparative analysis of savings conducted to assess the effects. Testing to find the optimal (most cost-effective) frequency can require a relatively large pilot but would pay handsome dividends if or when the program goes to full-scale deployment.
- Habits versus hardware? Are the savings estimated in the evaluation due to changes in habits, which may or may not continue? No longevity estimates are yet available for those. How can we get a handle on these? Are the savings in some part due to installation of measures, not covered by other DSM programs, that would be hard to undo and whose usage life are estimable? Until evaluations drill down to this level of understanding about the actions behind the savings, the reliability and persistence of the savings will remain elusive. *With* that understanding, it can be determined whether or when a behavior program meets the reliability standards for generating savings that other programs in a utility's resource acquisition DSM portfolio meet.

In order to tease out these influences, actions are necessary at several stages. Program planners and implementers need to enable evaluators to identify these different sources of the savings. For programs such as the comparative feedback ones used as examples in this paper, program planners must clearly incorporate random control elements in the program design, such as assignment of treatment customers to comparative groups as noted for testing messaging or treatment frequency; and the

[14] As example, see Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211(4481), 453-458.

implementers must maintain the integrity of the design, to enable statistically reliable estimation of savings. Evaluators need to then unveil and assess the various factors that contribute to those savings. This means going beyond making estimates of overall savings based on billing or interval data. The evaluations need to include follow-up surveys of participant to obtain additional information about the customers, their homes, and the actions that contributed to those savings; and quantitative analysis to parse those savings.

Conclusions and Recommendations

Design behavior programs to meet same standards for reliability as traditional DSM programs. In order for behavior programs to be ready for inclusion in a DSM portfolio, these need to pass muster with system and distribution planners when DSM is called upon as a preferred resource. As DSM integration into grid planning needs are considered, behavior programs need to be ready to answer the tough questions on reliability of these programs to meet resource reliability needs over an extended time period. These programs must be able to provide various types of data to fulfill the needs of demonstrable and persistent energy savings the grid or local distribution can rely on because power plants can be proven to deliver on demand whereas behavior savings need to prove its availability when needed.

Unveil what's behind the reported energy savings. Until we know what actions drive the savings, we won't have answers on persistence or how long program treatment needs to continue to maintain reliability. It is very important to understand when and where the savings are happening and how reliable these savings are to be relied on as a resource among other distributed energy resources. Unveiling customers' actual action-taking and assessing the persistence of those actions must become essential data elements for these programs to serve a purposeful role in DSM portfolio. For instance, are customers taking curtailment steps that develop into persistent habits, or will they revert to old habits once the treatment is withdrawn? Similarly, if these programs have effects on purchase decisions, are these effects permanent, or do they dissipate over time?

Develop more detailed program designs process. By their nature, behavior programs need an on-going "test-improve-repeat" approach. To keep the behavior programs and strategies refreshed, we need to build testing of alternate messaging and behavior change stimuli into the program design. Hence, addressing evaluation requirements within the program design process is essential.

Establish criteria for deciding whether a program is ready for inclusion in the DSM portfolio. We've identified a number of considerations, including: program motivation, demonstrated reliability and persistence of savings, and on-going and built-in evaluation process within program design for determining whether a behavior-based program should be included in the DSM portfolio. Each jurisdiction and customer market is unique though. Each utility needs to consider the characteristics of the regulatory environment and customers in making its own "checklist" to determine the appropriateness and readiness of including one or more behavior programs in the portfolio. These decisions especially become important when these programs are targeted to be included as part of DSM integration into system and distribution planning needs.