Widgets versus Actions: Measuring the Role of Behavior Change in DSM Programs

David Juri Freeman, Skumatz Economic Research Associates Inc., Superior, CO
Lisa A. Skumatz, Ph.D., Skumatz Economic Research Associates, Inc., Superior, CO

Abstract

The protocols for the evaluation of widget-based approaches to energy efficiency are well designed, documented, and for the most part, agreed upon in the industry. On the other hand, though many US and international organizations have touted the potential energy savings from behavior modification programs, the evaluation of these is still in its early stages. This research paper shares both a review of the current best practices in the measurement of behavior change found in the literature as well a hands-on case study measuring behavior modification impacts. The authors discuss the state of current practices in determining how an evaluator can identify what needs to be measured as well as the myriad of options to complete the measurement. The paper also addresses a host of other issues such as behavior retention, persistence, impacts compared to other approaches, and what is currently under-valued in behavior evaluation. Finally, the paper reviews recent projects completed by the authors, showing how the techniques discussed can be applied in the 'real world.'

Introduction

In today's energy efficiency marketplace, the number interventions that are strictly measure based is declining. While it is true that some interventions, for example improvements to building envelopes, may have little to no interaction with behavior, it is challenging to find many programs that have no behavioral component. The installation of an energy-efficient refrigerator, for example, will be impacted by the behavior related to cleaning the coils; a heating, ventilation and air conditioning (HVAC) system's optimal energy savings will only be realized through proper heating and cooling settings. Program managers realize this and are taking significant steps to develop behavioral interventions. Other efforts are underway to implement large-scale social marketing campaigns with little to no focus on widgets, and of course, there is a wide array of programs in between. Behavioral actions and choices, according to the American Council for an Energy Efficient Economy, represents lost energy savings of perhaps 30 percent or more with current technologies in the US alone (Earhardt-Martinez, 2009). Changes in how consumers behave can increase measure uptake, increase overall program participation, change the way consumers think about and utilize energy, impact maintenance and upkeep schedules, and help build stronger customer relationships. Yet, there are scores of ways to approach outreach, and not all of them are built the same. One potential approach that is gaining significant traction is the use of community based social marketing.

What is Community-Based Social Marketing (CBSM)?

Social marketing at its core is the use of traditional marketing techniques combined with sociological and psychological tools as a way to influence a target behavior. Classic social marketing campaign examples include efforts to curb teen drug and alcohol use, anti-obesity campaigns, recycling / solid waste, and of course, energy efficiency. A few of the aspects that make a social marketing campaign different than a traditional outreach campaign include the identification of barriers and motivations, targeting a specific sector as opposed to a more generic campaign, and the use of tools such as social norms, prompts, and feedback. Common and recent examples in energy efficiency projects that incorporate social marketing tools include real time feedback projects (uses feedback, norms, prompts, messaging) and changes to the way utility bills are designed to include comparisons to other households (uses norms, messaging, prompts). A time-tested example that is being improved upon through utility and community partnerships, social marketing, incentives and financing is the energy audit / measure installation (incentives, door to door outreach, social networks, norms, prompts). Despite the potential for all of these programs to significantly change behaviors and energy demand, the measurement and tracking of the behavior change programs lags far behind the tested evaluations for measure-based approaches.

Social Marketing Case Study Review

The authors undertook a comprehensive review of the literature related to social marketing and behavior change campaigns. The review included published reports, white papers, journal articles, conference proceedings, and web reviews. The search was not limited to energy efficiency only, but instead looked at all fields of behavioral research. The broad scope was used to determine what techniques are used to measure behavioral changes, find out what efforts are effective in achieving lasting behavioral change, and potentially, what the costs of those efforts are, regardless of the targeted behavior change.

The most common type of social marketing case studies found were health related (26%). Nearly one-fifth of all of the case studies researched were related to trash, recycling and/or composting, and 15% of the case studies reviewed were related to energy use and efficiency. We also found programs addressing water, transportation, and general environmental strategies. Some of the best-documented and well-known social marketing research projects have been conducted in the energy field. Luckily, unlike some of the other sectors addressed through social marketing, energy allows for easier reporting of impacts. It is relatively simpler to report on reductions in kWh compared to a reduction in teen binge drinking.

The review of the literature was illuminating not only in what was included in the published research, but also what had not been reported. We identified a number of common success elements in achieving behavior change; however, two very important areas of evaluation were missing in the published materials: cost and cost-effectiveness; and retention. If behavioral programs are to be considered seriously in the resource mix, these represent critical gaps in the literature.

1. No link between costs and impacts. While many reports included some information on impacts and some may have included information on total budgets, the number of reports that supplied information on the impacts on energy savings (or other impacts related to behavior change) per dollar spent was extremely limited. In some cases, this may be due to confounding factors (i.e., the behavior change was part of a wide portfolio of programs) while in others it is due to a lack of a control group, baseline measurements, or clear accounting of the costs of the outreach effort.
2. No Retention Results. An even more glaring omission in behavioral efforts is the issue of behavior retention. Estimates of widget-based effective useful lifetimes (EUL) have been incorporated into evaluation protocols for decades. However, there is extremely limited amount of information on how long a behavioral change lasts. The implications of the retention on determining the overall cost effectiveness of the efforts are significant. Unfortunately, even if first year annual savings estimates are available, it is not possible to develop reliable estimates of the benefit-cost ratio, nor is it possible to rely on long-term savings from programs that are not continually refreshed. For this reason, many

utilities assign retention values no higher than three years in most cases and a one year 'deemed' value for behavior retention is not uncommon.[1]

Findings on Energy Efficiency Impacts of Behavior Changes

The review of social marketing case studies included both commercial and residential programs. However, the vast majority of case studies reviewed for this study targeted the residential sector (over 80%). The case studies included energy audits, mass marketing campaigns coupled with door-to-door efforts, residential assistance programs, feedback programs, efforts at schools, public events, and public spaces (sports venues, government offices, fairs, etc.), and social media campaigns. While comparisons to a control group were limited to only two of the case studies reviewed, comparisons to simple pre-program baseline numbers were common (however, these baselines did not generally adjust for dynamic baseline conditions, weakening the transferability of the results).

The studies ranged from behavioral programs alone with no measure installations (a limited number of programs) to efficiency efforts in which the behavioral modification portion was one component of a larger program incorporating social marketing tools. The review found impacts from a 0% reduction in electric or gas use (and in one case an increase of residential energy demand!) ranging to a high estimate of 30% reduction in residential energy consumption (a multi-resource audit in Ontario using incentives, coupons, in-home demonstrations, and other tools).[2] Our research finds the average reported energy savings in the case studies was in the 5 - 15 % range. Similar results, ranging from 4%-12% savings, were found for the case of residential real-time pricing (feedback) pilots (Foster and Mazur-Stommen 2012). However, the case studies tend to highlight 'success stories,' and finding published information on typical program impacts is extremely challenging.

Not all behavior modification programs reported progress toward goals in reduction in use or demand. Some used metrics such as units sold or distributed (CFLs, ENERGY STAR® appliances), commitments or following through on commitments to behavioral modifications (turning off lights, using power strips, cold laundry, shorter showers, etc.), and still others measured success based on the number of audits completed. The costs of the programs varied significantly with smaller scale energy programs being in the tens of thousands, several in the low hundred thousand dollar range to million dollar range, up to a multi-year large-scale mass media and retrofit campaign – that also included specific door-to-door and social marketing initiatives - with a starting budget of over $23 million, and California's "Flex Your Power" (incorporating extensive social marketing) with a budget of $50 million per year.

Findings on Behavior Retention

Although very little research has been done on behavior change retention, there are three studies available that addressed the retention of educational messages and installation of low-cost energy efficiency measures delivered through energy education programs. The first, a multi-year study of the impacts of social marketing and door-to-door outreach on residential energy and recycling behaviors, indicates that personal interactions coupled with well planned social marketing strategies can significantly increase the adoption of energy saving behaviors and the impacts appear to have strong retention. One example is given in Section 4 of this paper. An earlier study by Harrigan and Gregory (1994), found 85%-90% of the savings from the education portion of a weatherization program was retained after three years. Publications of retention of behavior change are generally lacking; even well-funded multi-year statewide outreach programs have not examined the persistence issue.

---

[1] There is minimal research to support these values.

[2] A commercial program also reported impacts of slightly over 30%; however, this case included a behavioral component combined with building retrofits, and the behavior-only impacts were not provided.

Measuring Impacts

Measurement protocols for energy-related behavioral programs follow the same principles as many other types of evaluation work (Skumatz et al. 2010, Sergici and Faruqui, 2011, Sebold et al. 2007, and GAO 2009) and are not mysterious; the bottom line is that there are few substitutes for good up-front experimental design and random assignment.[3] An abbreviated summary of best practices principles follows. Although phrased in terms of residential programs, the principles extend to other programs and sectors.

- **Initial conditions**: Identify the goals of the program and the effects of interest (including a definition of what constitutes "participation" or "adoption"), and ensure that the effects can be seen to be caused by the program's intervention(s) (and not spurious factors). Assure that the program is administered to a group of participants that can be seen to represent the (ultimate) population of interest.
- **Experimental Design and Sampling**: Plan for a test and control group. Both the control and test groups should be large enough to support statistically valid and meaningful comparisons. Sample sizes supporting +/-5-10% at 90-95% confidence are preferred.[4] The control group should be as similar as possible to the test subjects (and both groups should be as similar as possible to the ultimate group that will be eligible for the program to maximize transferability of results). Pre-post measurement of the test group is not best practice. Pre-post alone is vulnerable to seasonal differences, and other factors; control groups allow easy and reliable netting out of these variations. The control group sorts out impacts from effects beyond the program (e.g. nationwide ads from EPA or others, etc.), and serves as a dynamic baseline against which the effects can be measured to provide net impacts. The premier experimental design is random assignment of eligible customers into the test and control groups (Skumatz et al. 2010). Random assignment also helps to eliminate self-selection bias. Other approaches that have been taken include use of "similar" counties or cities, neighboring / similar states, etc.[5] Controlling for other factors from these "similar" control groups can be attempted through corrections with statistical models, but random assignment is much more straightforward and reliable.
- **Measurement**: Evaluation methods need to be clearly laid out before any data collection is conducted. When evaluation is concluded, all limitations of methods and results need to be clearly identified. In addition, the evaluation should include an assessment of the associated uncertainty, and the evaluation should identify the way in which the impact(s) will be measured. For energy behaviors and energy savings, there are several main approaches:
  - **Metering**: If the project (and budget) allows, metering the equipment affected by the desired behaviors over the course of the experiment provide direct and reliable information on the behavior change and its energy impacts. If inexpensive metering approaches can be identified, the metering should be installed in large[6], random (or representative) samples of

[3] This is the approach illustrated in the case study that is described later in this paper.

[4] Assuming large populations, these requirements tend to require sample sizes of 68, 96, 270, or 380 observations (for each subgroup to be examined; see information on the frequently-ignored impacts of CV variations on sample in forthcoming Baker, 2012), with higher numbers preferred. Greater specificity on sample size depends on the degree to which the measurement needs to address Type 1 error, Type 2 error, one or two sided hypothesis testing, single or repeated measures experiments, etc. However, a surprising number of social marketing programs have measured the impacts based on sample sizes substantially less than 100 (often 30), and only pre-post and not control group measurements, to the detriment of confidence in the results.

[5] As widespread education campaigns affecting both target and non-target audiences become more common, finding a baseline to measure against is more difficult - it is hard to uncover a population with a "zero" behavior baseline.

[6] Obviously the principle is that large samples reduce the variance and help detect significant differences between the groups.

the test and control groups. If not, metering should be installed in smaller, strategic samples that support generalization to larger samples (Vine 2012).

- **Utility bills and impact evaluation**: Preferred data for this option include monthly energy usage (and billing cycles/meter reading dates and possibly tariffs) for sampled treatment and control customers.
- **Surveys and reported behaviors:** The researcher needs to identify the relative appropriateness of phone, in-person, web, or other types of surveys. Research should include well-crafted / tested question methods, for example, asking about behaviors undertaken in specific time frames, rather than "general" habits,[7] and other preferred survey approaches.[8] Again, control groups are highly recommended to provide "baseline" behaviors.
- **Demographic Information**: Gathering information on number of occupants, socio-demographics, appliance data, occupancy (move-ins, etc.), weather data, and other information can help in developing statistical models that control for these sources of variations in results when conducting impact evaluation work or other comparisons.

- **Impacts and Analysis**: The basic preferred analysis approach is a comparison of means between treatment and control group, using either one pre/post period, or periodic measurements. The appropriate tests for statistically significant differences are performed to identify impacts from the program. Multiple measurements over the course of the project / pilot provides advantages in efficiency and variance reduction (due to the correlation between measurements at different time points), and thus, greater confidence in the results. Analysis of the data up-front is valuable (monthly comparisons, plotting data, and conducting comparisons of "features" between test and control groups to ensure comparability). Impact evaluation work using statistical models and the energy data can provide reliable estimates of these means. This may employ one of several methods for estimating impacts:
  - Measurement and Verification (M&V): using metering or estimating key parameters from a random sample (or all) of the participants and control group and applying to all members of the group.
  - Statistical Analyses: applying statistical regression models[9] to utility billing[10] or metering data of sampled program participants, including approaches like differences of means / ANOVA, difference in differences[11] and panel data regression analysis[12], and other methods

[7] For example, 'did you use the power strip yesterday', or 'how many of your last two laundry loads used cold water', rather than 'do you use power strips' or 'do you use cold water for your laundry'.

[8] To help increase confidence in the survey reports, the researcher might also research the "say / do" gap and run scenarios on the range, conduct a sample of on-sites to confirm some behaviors that might be easily observed (current laundry temperature settings, etc.), and benchmark against a sample of billing data, where possible.

[9] Econometric texts can address the issues involved in regression modeling including model misspecification, measurement error, correlations, etc.

[10] Billing analysis using weather-normalized consumption data provided by the utility commonly is used to estimate gross savings. Billing analysis requires consistent residency for two or more years, so one year of pre-program data can be compared with one year of post-program data. Billing analysis may be used to estimate gross savings of education programs combining low-cost measures and behavior modification. However, as billing data are inherently too "noisy," gross savings less than 10% of pre-consumption levels are hard to detect.

[11] This involves netting out mean differences between treatment and control groups in the pre-treatment period from the mean differences between treatment and control groups in the post treatment period. If the difference in differences of the mean values is statistically significant, then the treatment is found to yield an observable effect in the usages of the treatment customers (Sergici and Faruqui, 2011).

[12] Advantages of this approach include the possibility to increase the efficiency and precision of the estimate using repeated measures on each program participant and to account for time invariant unobservable variables that would otherwise lead to biased estimates. Modeling approaches include fixed effects or random effects models (Sergici and Faruqui, 2011).

provide reliable estimates of impacts. Cross section and time series approaches are valid. There is an extensive literature on statistical, or statistical / engineering adjusted models.
    - Surveys and Self-Reporting[13]: surveying certain populations to gather information regarding knowledge or behavior to estimate the savings-related changes from behavioral / educational / social marketing programs, and analyzing for statistical differences in the adoption of the behavior. Assuming energy savings and kWh are the key impacts of interest, an additional step involves identifying an estimated or deemed value for the savings "per adopted behavior" – which may be the best information on overall savings available from this method.
- **Retention**: Measurement over the course of a full year for behavioral programs is preferred, to help account for seasonal effects. However, we would recommend working to put in place a measurement protocol that follows beyond that period to test for retention of the effects – which is a very important uncertainty component of behavioral programs. The length of time a widget-based program will have an effect is usually directly related to the retention of the measure in the home or business;[14] the retention of an adopted behavior change lasts only as long as the behavior remains changed. The retention (and impact) issue may suffer from total cessation of the behavior, occasional retention by an actor, or retention by only some occupants of the home. Conversely, some behaviors may form new habits and remain in place for a lifetime. All of these possible changes – and more – will have an effect on the lifetime (and level of) of the estimated savings from the program.
- **Deviations and Alternatives**:[15] Although random assignment is the "gold standard", the world (and budget) does not always allow for this design – particularly if large-scale broadcast media are used, and potential participants cannot be excluded (see Sebold, et.al. 2001). Other options include:
    - Quasi-experimental comparison groups,[16]
    - Statistical analysis of observational data,[17] or
    - In-depth case studies or other approaches.

Case Study

In 2009, the authors conducted a social marketing experiment on 1,600 households in suburban Colorado. The project was tasked with delivering conservation – in the form of energy and recycling. However, as part of the project, we built in an experimental design that would support detailed and defensible evaluation of the results – including the two areas of cost-effectiveness and retention. We conducted detailed cost recording and 12 months of post-outreach measurement to measure the impacts and cost effectiveness of using social marketing to change behaviors. We placed the homes into three groups, each designed to receive one of three levels of outreach and social marketing interventions: (1) control group, with only a standard (limited) outreach (e.g., an annual mailer to the homeowner

[13] Self reported data are often augmented with site visits and selected metering (e.g., hours of use).There are many texts that address the issues in survey design, question development, bias reduction, etc. This is not addressed here.
[14] And can be estimated using accepted and adopted median measure lifetimes or EULs (Skumatz et al. 2010).
[15] Note that the GAO report also stated that improvements to any evaluation can be achieved by: collecting additional data, targeting comparisons, and gathering a diverse body of evidence (GAO 2009).
[16] These resemble randomized experimental design, but the groups are not randomly selected and are selected from un-served members. This might include groups denied participation when a program is full or those that will participate in the next period, etc. According to one report (GAO, 2009), the approach requires statistical analysis to establish groups' equivalence at baseline, and potentially, specialized statistical modeling in some examples, such as regression discontinuity analysis (GAO 2009)
[17] This approach requires observing and collecting data prior to the intervention, and after the intervention.

highlighting what can be recycled and how they can recycle and city wide information on energy efficiency); (2) social marketing (including prompts commitments, norms, barrier identification, etc.) delivered through door hangers and mail; and (3) the same social marketing outreach as the second group, plus delivery of materials via a door-to-door outreach campaign.

Baseline Measurement: Baseline measurement is necessary to measure incremental impacts attributable to the project's interventions, and these impacts are essential to computing the relative costs-per-impact for the two test routes. Different methods were used to establish the energy and the trash/recycling baselines. Establishing reliable trash and recycling baselines were important because, despite requests (the utility was not a project sponsor / participant), we were unable to access energy bills. We recognized that tracking the (relatively more easily monitored) recycling behaviors – and particularly, their retention, could serve as a proxy for the potential pattern of retention of energy behaviors.

- *Energy Baseline:* Collecting metered data on energy use for the routes was not possible for the project. Instead, proxies for energy behavior were established and collected through surveying, focus groups, and one-on-one household visits. The authors used metrics such as reported average monthly energy bills[18] (for pre/post comparisons), the number of CFLs in each house, the frequency / likelihood that respondents undertook energy-efficient behaviors such as purchasing ENERGY STAR® appliances, reduced automobile idling time, used low-flow shower heads or limited shower times, and others. We also gathered information on energy knowledge items, which we compared between routes and as pre/post comparisons.
- *Trash and recycling baseline:* A baseline of set-outs, trash collection, recycling collection, participation, diversion rates, and subscription levels was established as a starting point for before, during, and after project comparisons. We collected on-going weekly data on recycling participation and tonnage, as well as other periodic measurement of household-level trash / recycling activity.

Comparisons of Committed Actions and Energy Behavior Changes

Our analysis shows that the committed activities would save 360 MTCO2E per year: In the two test routes, 12.5% of all households made the commitment to undertake energy saving and recycling actions (none of the households in the control group made a commitment, as they weren't asked / notified / included – therefore, these are "gross" estimates, and not "net" estimates). When residents took the commitments, they were asked whether they were committing individually or for their entire household. When all household members are included, more than 500 people committed to taking green actions. Overall, there were 2,300 committed green actions. Figure 1 displays the distribution of the committed actions.

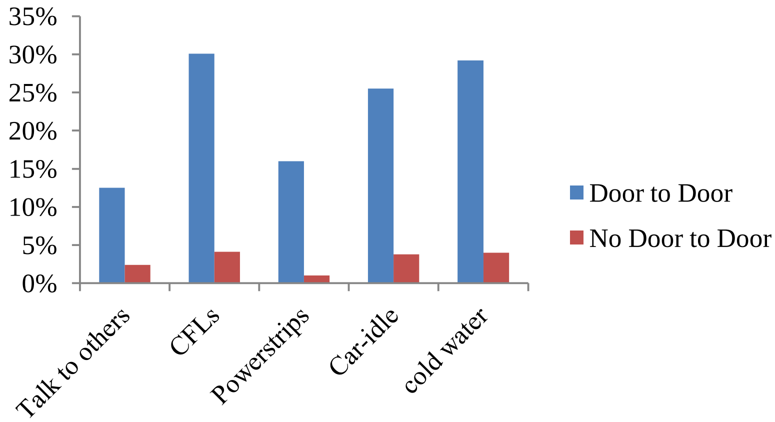
Figure 1: Household Committed Green Actions (n=214)

Install 1 CFL	Use power strip to turn off electronics	Turn off car when idling	Use cold water for laundry	Talk to one neighbor	7 pounds more recycling	Recycle all paper and cardboard	Use one re-useable bag when shopping
69.6%	34.4%	56.0%	68.8%	28.0%	52.8%	83.2%	63.2%

[18] We got perhaps 60% response on the question, with a range of responses (and a few humorous responses), We could not conduct any tests for accuracy, and it was not clear households would recall the values accurately. The question was included as a test to see if it might be useful. This, and all the succeeding metrics suffer from potential problems from recall, as do all these types of survey-derived data collection..

In the door-to-door outreach group (Group 3), 42% of all households made a commitment to increase their energy efficiency, compared to just 4% of the partial treatment homes (Group 2) – 10 times better penetration of these commitments (Figure 2). Figure 2 also displays the percentage of total households in each test group committed to various energy savings actions. The most common action was a commitment to install a CFL followed by a commitment to wash one load of laundry per week with cold water. The least popular action was *"Talk to a neighbor about energy efficiency or recycling"*.

Figure 2: Commitments to Energy Actions



A post-outreach survey was used to determine variations in impacts between groups as well as to test residents on message recall. Figure 3 displays message recall and familiarity with the campaign. While there appeared to be some very minor spillover (7% of the control group had heard 'something' about the campaign but were unfamiliar with the campaign), the differences on recall between the door-to-door and partial treatment groups were stark. None of the survey respondents in the partial treatment route reported they were 'very familiar' with the campaign, while nearly one in six respondents in the full outreach group reported they were very familiar with it. The same differences between routes are echoed in their reported energy actions in the post-outreach survey (Figure 4).

Figure 3. Survey Responses to Familiarity with the Social Marketing Campaign

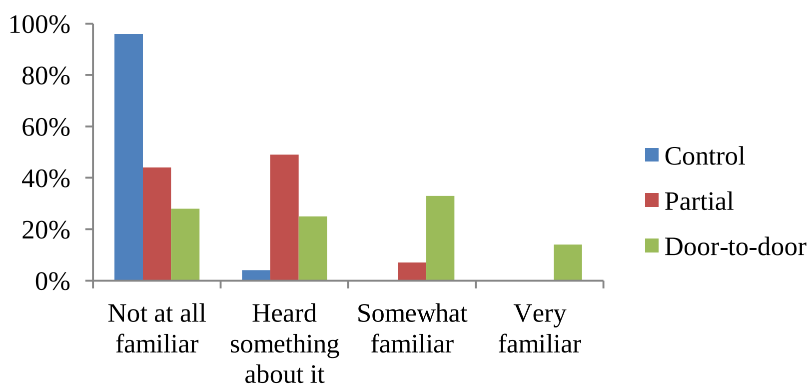


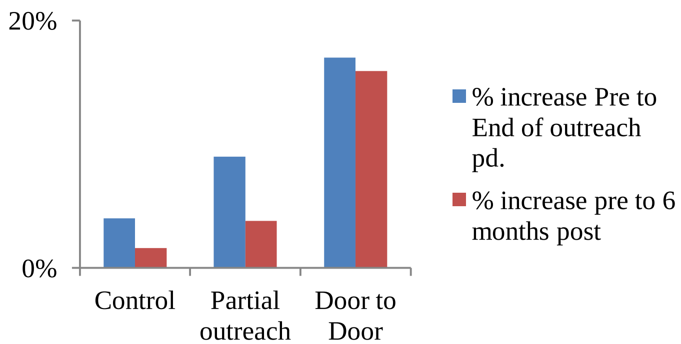
Figure 4: Energy Behaviors

	Control	Partial outreach	Door-to-door outreach
Number of last 2 loads of laundry that were rinsed in cold water	n/a	1.48	1.76 (19% more than partial treatment)
Turned off power strip yesterday (% yes)	11.3%	17.6% (56% more than control group)	20.5% (81% more than control group)
Adjusted thermostat up one degree in summer and / or down one degree in winter	41.9%	54.5 % (30% more than control)	55.3% (32% more than control)
Installed caulking (% yes)	n/a	11.8%	36.8% (312% more than partial treatment)

Costs, Impacts, and Retention

The surveys provided self-report information on adoption of energy behaviors[19], and these data were combined with the measured data gathered on impacts on weekly recycling tonnages (using net comparisons between impacts and control groups). Both energy and recycling impacts were translated into metric tons of carbon equivalents (MTCE). [20] Costs for outreach dollars spent per increase in MTCE greenhouse gas (GHG) reductions were calculated. Group 3, the full social marketing campaign using door-to-door outreach, saw an almost 20% increase in recycling: this was nearly double the increase the social marketing campaign delivered without the door-to-door approach and almost 5 times more than the traditional outreach campaign. The energy impacts, and differences between routes, were shown in Figure 4. We also conducted some work to reflect retention (and the retention tracking is continuing). When we went back 6 months later to see if residents were continuing their recycling habits, the door-to-door group was still recycling significantly more than before the outreach while the non door-to-door regressed closer to their pre-outreach levels. The comparison of recycling and retention at the end of the social marketing outreach campaigns and six month retention figures are shown in Figure 5. Interestingly, the during-program recycling increases for the two groups were in or near the range of the percentages found for energy savings from the literature review.

Figure 5: Recycling and Retention Results by Group



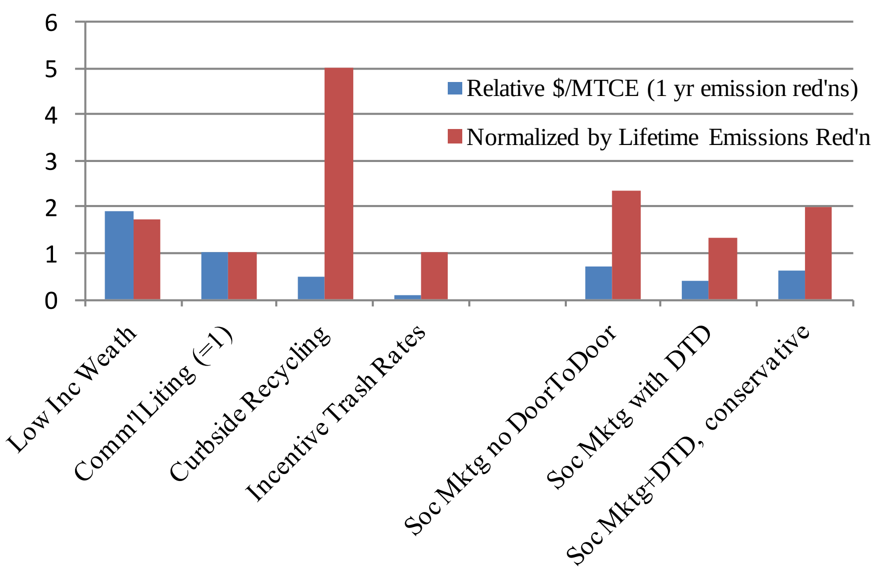
[19] Independent measurement of these activities was not possible. The surveys included questions about specific behaviors in a specific time period (yesterday, this week) to provide better quality data than "general" behaviors-type answers. We also considered and incorporated a range of values for the "say / do" gap.

[20] We calculated the kWh and GHG impacts of the following net energy behaviors (net beyond those reported by the control group): changed temperature setting one degree, used a power strip for electronics, installed caulking, and washed laundry using cold water. We omitted numerous other energy (CFLs, etc.) and transportation behaviors (idling) also reported in the surveys. Our computations found the energy savings impacts on MTCE were multiple times those from the recycling behaviors.

The door-to-door campaign was several times more expensive on a per household basis than the mail and door/cart hanger approaches, but the cost per MTCE from the program was less than half as expensive for the door-to-door social marketing campaign as the mail / door hanger campaign. However, the most useful cost comparisons put the cost for social marketing in the context of other alternatives upon which the funds might be spent.

In several previous papers, the authors have computed the cost per metric ton of carbon equivalent (MTCE) emissions avoided from several standard widget-based programs and generation alternatives (Skumatz and Freeman 2011, Freeman and Skumatz 2010, Skumatz 2009). The results for a number of energy efficiency and recycling alternatives are shown in Figure 6, with the results normalized to show multiples of the cost per MTCE (one-year, normalized to cost per MTCE for commercial lighting=1). Taking one additional step, we can control for lifetimes of the measures, programs, or facilities. With this computation, we find the costs per MTCE from this pilot social marketing program would be in the range of other widget-based energy efficiency programs.[21] These estimates would be revised downward for full-scale social marketing programs, and upward if the social marketing measures lasted a shorter time. Although more research is needed[22], we find that the cost per MTCE and potential for these types of social marketing programs – even relatively expensive small-scale pilot versions – are not highly unfavorable relative to some traditional measure-based options (see Figure 6), and preliminary work shows the results are also in the range of some generation alternatives.[23] Full scale implementations may show improved performance.

Figure 6: Relative Costs per MTCE Emissions Avoided by Strategy
(Normalized to Commercial Lighting =1; one set of scenario assumptions)



[21] Assuming a 10 year life for the commercial lighting, 11 years for weatherization, and 3 years for the behavioral measures. Changes in these assumptions would lead to revisions in these illustrative computations. According to research by the authors, lifetimes of 3 years are assigned by several utilities; certainly, as this paper points out, these estimates require further investigation.
[22] The SERA research presented here and cited in this paragraph was internally funded.
[23] Preliminary computations show results may be in the range of cost per MTCE from biomass and wind alternatives, for example. Results can vary based on a wide variety of assumptions, including generation mix.

Summary and Conclusions

There are legitimate concerns about behavioral programs and social marketing efforts.

- The savings have been well-measured (including control groups) in only a few cases, and programs are all distinct, potentially leading to different savings values with variations in design, target, etc.
- The programs have mostly been pilot in nature; full-scale implementation results may lead to different savings and costs results.
- Costs, cost-effectiveness, and retention are rarely measured.

However, behavioral programs have several major advantages when compared with traditional widget-based programs:

- They can have significant impacts on energy use (individual pilots and social marketing programs commonly show impacts on the order of 5-15% savings (Green and Skumatz, 2000, Skumatz et. al., 2010) – which reflects an enormous potential realized by few purely widget-based programs. ACEEE estimates that behavioral actions and choices represent lost energy savings of perhaps 30 percent or more with current technologies in the US alone (Earhardt-Martinez, 2009).
- They can be implemented quickly, with widespread adoption in a matter of weeks to months.
- They do not require programmatic purchases, delivery, or installation of equipment, intrusions into homes, and other efforts. The avoidance of these types of barriers may add additional value when considering program alternatives.
- The retention from social marketing is still a question. However, preliminary results indicate door-to-door methods lead to stronger retention than outreach of the same materials by mail, making the results potentially stronger than those already known for traditional outreach programs.

Integrated energy efficiency plans need data on energy, cost, and the number of years that the energy savings will be "accounted for." Behavioral and social marketing programs have generally been small parts of these plans, at best, and may be significantly undervalued in portfolios if indicative research bears out. The two most contested issues that we identified in this and previous studies (Green and Skumatz 2000, Skumatz et. al. 2010) remain cost-effectiveness and retention. More research on these questions is essential if behavioral programs and social marketing / outreach programs are to be a more integrated and reliable part of the energy efficiency portfolio. Behavioral programs have tremendous potential in addressing these issues: for example, our research on relative costs show results in the range of other widget programs, and retention from at least some types of programs are promising.

References

Baker, Andie, 2012, "Sample Size Selection in Energy Efficiency Research and Evaluation – The Use and Abuse of the Coefficient of Variation", Proceedings of the American Council for an Energy Efficiency Economy (ACEEE) Summer Study on Buildings, Asilomar, CA, August (forthcoming).

Earhardt-Martinez, Karen, 2009. "*Testimony before the United States House Committee on Science and Technology"*, Subcommittee on Energy and Environment. April 28.

Foster, Ben, and Susan Mazur-Stommen, 2012. "*Results from Recent Real-Time Feedback Studies"* ACEEE Report Number B122, American Council for an Energy Efficient Economy (ACEEE), Washington, DC, February.

Freeman, D. Juri and Lisa A. Skumatz, Ph.D., 2010."A kWh is not just a kWh: Comparing Various Energy Efficiency Programs in Terms of GHG, Job Impacts, and Policy Achievements (NEBs and Beyond)", Proceedings of the American Council on Energy Efficiency (ACEEE) Summer Study on Buildings. Asilomar, CA, August.

Green, John, and Lisa A. Skumatz, Ph.D., 2000. "Evaluating the Impacts of Education / Outreach Programs – Lessons on Impacts, Methods, and Optimal Education", Proceedings of the American Council on Energy Efficiency (ACEEE) Summer Study on Buildings, Asilomar, CA, August.

Harrigan, Merilee, and Judith M. Gregory, 1994. "Documenting Energy Savings Enhancements from Energy Education Components for a Low Income Weatherization Program", Proceedings of the 1994 American Council on Energy Efficiency (ACEEE) Summer Study on Buildings, Asilomar, CA, August.

Sebold, Fred, Alan Fields, Lisa A. Skumatz, Shel Feldman, Miriam Goldberg, Ken Keating, and Jane Peters. 2001. "*A Framework for Planning and Assessing Publically Funded Energy Efficiency"*, Study ID PG&E-SW040, Report prepared for Pacific Gas and Electric Company, San Francisco, CA.

Sergici, Sanem, Ph.D. and Ahmad Faruqui, Ph.D., 2011, "Measurement and Verification Principles for Behavior-Based Efficiency Programs", The Brattle Group, Prepared for Opower, May.

Skumatz, Lisa A., Ph.D., 2009. "Do Energy Efficiency Strategies Outperform Recycling in GHG Mitigation and Job Creation?", International Energy Program Evaluation Conference. Portland, OR, June.

Skumatz, Lisa A., Ph.D., and Juri Freeman, 2011. "Outreach, Social Marketing, and Behavioral Programs: Results on Retention, Cost-Effectiveness, and "Making an Impact", Presentation to the American Council for an Energy Efficiency Economy (ACEEE), Behavioral Energy Climate Change (BECC) Conference, Washington DC, December.

Skumatz, Lisa A., Ph.D., Dr. M. Sami Khawaja, Jane Colby, 2010. *"Lessons Learned and Next Steps in Energy Efficiency Measurement and Attribution: Energy Savings, Net to Gross, Non-Energy Benefits, and Persistence on Energy Efficiency Behavior"*, Report prepared for California Institute for Energy and Environment (CIEE), Berkeley, CA, November.

United States Government Accountability Office, 2009. "*Program Evaluation: A Variety of Rigorous Methods Can Help Identify Effective Interventions",* November.

Vine, E., 2012, Communication with the author, April 2012.