

# **How can we better evaluate the Success of Marketing Outreach Efforts?**

*Anne Dougherty, Opinion Dynamics Corporation, Oakland, CA*

*Brad Kates, Opinion Dynamics Corporation, Waltham Mass*

*Jennifer Mitchell-Jackson, Opinion Dynamics Corporation, Oakland, CA*

## **ABSTRACT**

As marketing and outreach efforts as used in more and more jurisdictions to help promote energy efficiency, regulators are asking whether marketing and outreach program's can be effectively quantified. The Opinion Dynamics (ODC) evaluation team is currently evaluating one of the largest social marketing campaigns in the energy industry, Flex Your Power, and has specific insights into how marketing and outreach programs can be evaluated.

In this paper, we discuss how adjacent industries, such as public health and social marketing, can serve as examples to energy efficiency evaluators. Namely, we will discuss the use of efficacy and effectiveness studies and how these two evaluation approaches may be combined to lend greater depth and insight into marketing and outreach evaluations. To help illustrate the use of these types of evaluations, our team will draw on innovative methods that we have used in our evaluation of the Flex Your Power program. Here, we will discuss the used of traditional effectiveness methods such as Pre/Post studies and then outline different efficacy and qualitative studies we are using that help to illuminate the success of the Flex Your Power program along the behavior change continuum. We will discuss Structural Equation Modeling, Verified Reach Assessments, Ethnographic Research, and Social Network Analysis. We argue that future evaluation efforts of marketing and outreach programs must include methods that evaluate programs from a 360 degree perspective and account for multiple potential successes along the behavior change continuum.

## **Introduction**

As the energy efficiency world has matured, the industry continues to explore ways to increase the penetration of energy efficient equipment. In the past, engineering concepts dominated, and incentives and rebates were the only tools in a program's toolbox. This approach has met with limited and niche levels of success--rebates and incentives do not speak to everyone. To overcome the limitations of economic incentives, the industry has recently begun to embrace the "four P's of marketing": product, price, place and promotion, rather than focusing only on price. With this recent change in philosophy new challenges in assessing the success of an energy efficiency effort emerge. Within the traditional economics-driven paradigm such as a strict rebate program, success may be measured by determining how many pieces of equipment were installed and working. Thus, it is relatively easy to determine program success because rebates are handed out and the program implementers know the names of the people who purchased the equipment. This is not true when program implementers begin to take a more comprehensive approach to motivating behavior change and draw on other less measurable motivations and drivers.

For marketing and outreach efforts, implementers are taking a more broad-based approach to behavior change. Because of this, we cannot easily attribute the purchase of an energy efficient piece of equipment to the marketing effort like we could from a rebate program effort. This paper provides some innovative approaches to help quantify the success of such marketing and outreach efforts.<sup>1</sup> Here, we pose three key questions: (1) What can the energy evaluation community learn from the evaluation of other social marketing and communications efforts?; (2) What are the most appropriate methods and techniques that may be applied to marketing and outreach evaluations?; and (3) how can utilities and evaluators use these methods to gain a better picture of program results?

Currently, there is a great deal of debate among evaluators on how we should measure the ever-increasing number of marketing and outreach campaigns. In California, protocols were put into place outlining the level of rigor to be used for marketing and outreach program evaluations (TecMarket Works Team, 2006) and may be the most rigorous evaluation criteria for marketing and outreach campaign ever mandated by a regulatory body. While these protocols are intended to hold marketing and outreach programs accountable to their outcomes, they are the subject of a great deal of debate. Namely, an evaluator's ability to accurately and consistently measure program *net effects* using standard participant recall is still in development, and the industry has yet to outline the breadth of potential program *effects*, how we define those effects, and perhaps the biggest challenge, how we measure them. Here, we focus entirely on how to evaluation effects, but future papers and greater attention should consider how these definitions are drawn and what metrics may be identified to ensure that programs are evaluated from the 360 degree perspective.

### **What can the energy industry learn from other marketing and communications efforts?**

The field of public health, which relies heavily on social marketing principles, utilizes marketing and outreach efforts to generate behavior change around a number of key issues such as obesity, HIV and STI prevention, and anti-smoking efforts. Such marketing and outreach campaigns as we know them have been applied in this field for over half a century and may be traced back to the publicly funded Work Projects Administration (WPA), who used print and radio formats to promote personal hygiene and public safety during wartime. However, such efforts for the energy efficiency industry, independent of specific resource acquisition and demand-response programs, are still in their infancy. Given their long-standing history in social marketing, energy efficiency evaluators should look to the public health sector for guidance on program implementation and evaluation.

**Efficacy vs. Effectiveness Studies.** The public health sector employs two primary types of evaluations for health communications and marketing campaigns: efficacy and

---

<sup>1</sup> We refer to marketing and outreach efforts throughout the paper. By "marketing and outreach" we mean any broad-based marketing effort that is not directly attributable to a given energy efficiency program effort that involves rebates or other incentives. In California, the Flex Your Power marketing effort is a perfect example, as is the EnergySmart marketing effort in New York.

effectiveness evaluations. As outlined in recent presentation by W. Douglas Evans, Ph.D., at the American Evaluation Associations 2008 National Conference in Denver, these two types of evaluations use similar measurements, but differ in their approach. Efficacy studies control exposure to messaging in their experimental design, while effectiveness studies measure exposure through self-reported recall or environmental measures. In public health evaluations, effectiveness studies tend to be the industry norm. Two key examples of these efforts are the American Legacy Foundation's Truth® campaign, which is an anti-tobacco campaign aimed at teenagers, and the National Drug Control Policy's evaluation of their anti-drug campaigns aimed at changing how teenagers and parents discuss drug use to one another and among their peers (Evans, Uhrig, Davis, & McCormack).

In energy efficiency evaluations, we tend to rely on effectiveness studies to determine the success of marketing and outreach programs. Namely, evaluators use random digit dialing survey efforts to generate metrics of self-report. In certain states, such as California, algorithms are applied to self-reported figures to tease out net to gross effects of marketing and outreach programs. Such algorithms, while often well thought out, do provide evaluators with a proxy for program net effects, but may not be the most rigorous approach to assessing the true impact of marketing and outreach programs when used in isolation. While effectiveness testing can provide the foundation for evaluation efforts, it leaves many questions unanswered and does not provide in-depth assessments of how a population and its segments move along the behavior change continuum—this is where efficacy approaches can provide great value.

**Efficacy Approaches In Social Marketing.** In social marketing, efficacy studies are typically used to test marketing and outreach campaigns during creative development to better align outreach messaging with their target audience(s). Effectiveness studies are more commonly used in program tracking and evaluation efforts. In their benchmarks for a successful social marketing campaign, the National Social Marketing Center emphasizes the use of efficacy studies in program scoping and planning, and effectiveness studies to track and monitor marketing and outreach effects over time (National Social Marketing Centre 2007). While on the surface these two types of studies appear to have decidedly different objectives throughout a program's life cycle, social marketers agree that these two studies should speak to one another and rely on the same well-defined and *measurable* metrics of success.

**A New Approach.** In the energy and public health industries, efficacy studies are generally seen as the role of marketing and outreach creative teams, and are rarely employed as a tool to measure program *effects*. However, energy efficiency evaluators stand to benefit from more innovative techniques that seek to join both efficacy and effectiveness techniques. In the final section of this paper, we discuss the potential application of multiple methods that combine efficacy, effectiveness, and qualitative data to provide a more illustrative and multidimensional portrait of marketing and outreach success.

**What are the most appropriate methods and techniques that may be applied to energy marketing and outreach evaluations?**

Unlike standard resource acquisition programs, marketing and outreach programs are particularly difficult to evaluate as they typically require self-reported awareness, attitudes, and behavior changes to quantify program effects through pre/post testing. As touched on earlier, this type of effectiveness study has been the standard but it is riddled with potential biases. Below, we outline how the Opinion Dynamics evaluation team has drawn on a wide range analysis tools to assess the effects of Californian's Statewide Marketing and Outreach program, Flex Your Power.

The Opinion Dynamics evaluation team utilized these methods to create a full picture of the Flex Your Power program's effects. The effort combines both efficacy and effectiveness studies into a single evaluation plan, and uses each to help adjust and temper the findings of the other. In addition, the Opinion Dynamics evaluation team has utilized a consistent and well-tested battery of questions designed to measure attitudes, knowledge, actions, and beliefs across all of our efforts. Through this attention to detail and the combined use of traditional pre/post analysis methods and more innovative techniques and approaches, the Opinion Dynamics evaluation team has developed a marketing and outreach-specific toolbox to employ in future evaluation efforts.

**Efficacy Approach: Verified Reach Assessment.** Efficacy studies, when combined with traditional pre/post analysis, can add depth to traditional evaluation approaches. The ODC evaluation team sought to conduct our own efficacy study by tracking panelists' exposure to the Flex Your Power program. Drawing on a panel of 300 participants in the Los Angeles DMA, the ODC team was able to verify panelists' exposure to messaging using Integrated Media Measurement's cell phone technology. This technology records all audible media that each panelist is exposed to, both in and out of his or her home, and thus can capture each panelist's exposure to TV and radio media (the Flex Your Power programs' most prevalent outreach channels). For the first time ever, we utilized this new technological approach to determine actual exposure, as opposed to only utilizing self-reported information. In this way, the verified reach assessment can give us a clear and definitive indication of one's exposure and compare that directly their actions. In this sense, panel provides a built in control group.

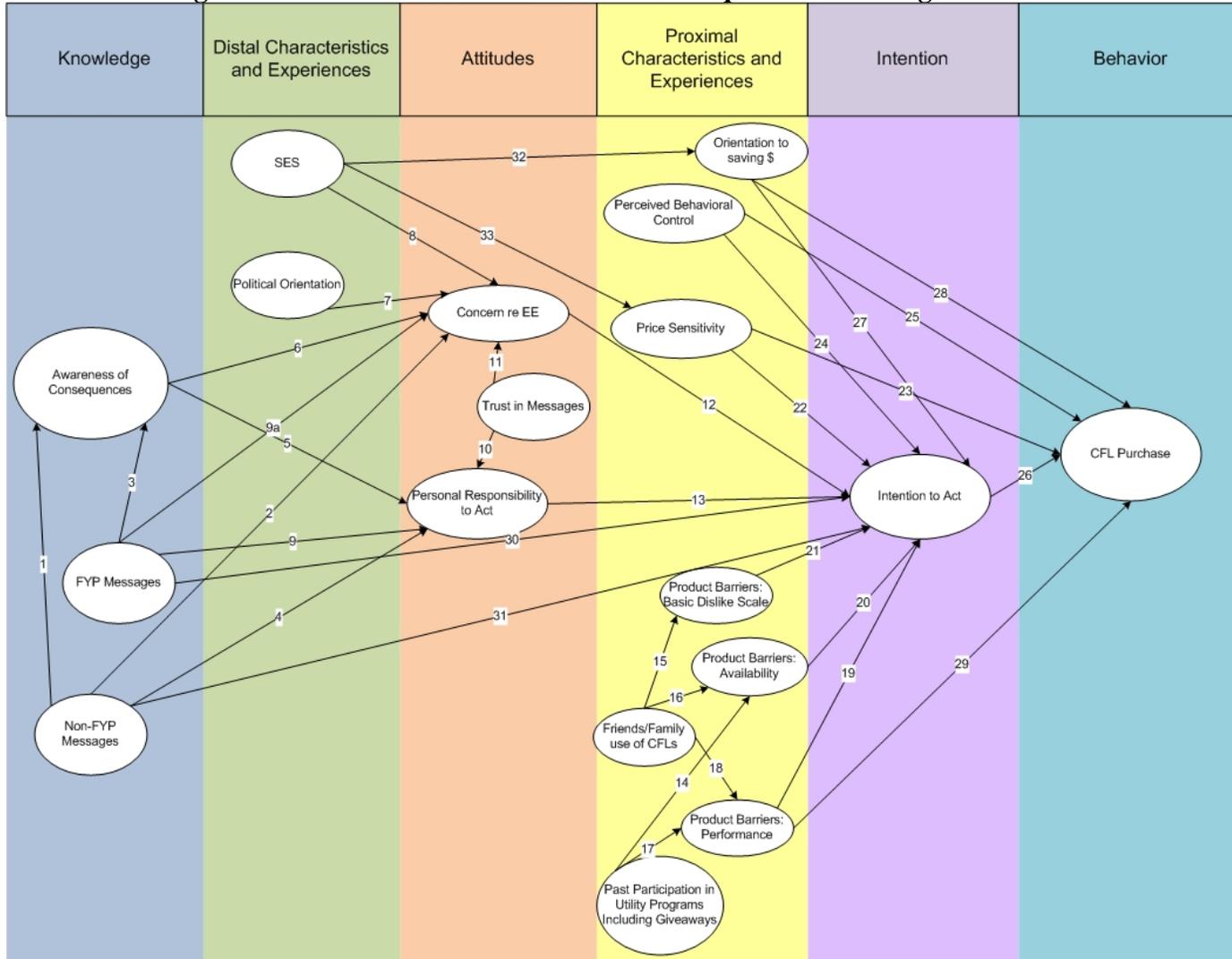
To conduct this analysis, the ODC team sought to explore a number of key questions: (1) What is the *actual* reach and frequency of marketing?; (2) How aware are panelists of specific marketing and outreach mass media messaging accounting for their relative exposure?; (3) How does this awareness change with increased exposure to messaging over time?; (4) Are consumers associating the FYP brand with its messaging and do they take the targeted behavioral changes?; (4) Are particular Marketing and Outreach messaging failing to resonate with particular sectors, sub-sectors, and segments while accounting for their relative exposure?; and (5) what is the rate of false negatives and false positives on program awareness claimed by self reported information when compared to actual exposure?

By answering these questions through an efficacy study approach, the ODC team has been able to develop adjustment factors for all of our tracking study recall, thus allowing our team to more effectively account for recall biases. In addition, this technology tells us what statistical modeling cannot: the effects of known, documented exposure on awareness and behavior change.

**Efficacy Approach: Structural Equation Modeling.** As another cog of our triangulation method, the Opinion Dynamics team utilized Structural Equation Modeling to assess the impact of program messaging on the CFL purchase behaviors of individuals exposed to program messaging. Structural equation modeling (SEM) controls for program influences through statistical modeling, rather than through controlling exposure. Unlike pre/post analysis, SEM provides a broad picture of influences that extend beyond exposure to messaging, thus laying clear more elaborate causal pathways to action.

The data collection for SEM includes a web survey of approximately 1,000 respondents. Here, we chose to focus on the most basic of purchase decisions in the energy efficiency portfolio, the CFL, which lends insight into how we might begin to assess more complex and economically taxing purchase decisions, such as an energy efficient heating or cooling system. Using this statistical technique, the Opinion Dynamics team was able to measure the relative influence of a range of factors on participants' efficiency behaviors. These factors include all those cited in the behavior change literature, as well as a few additional measurements: other energy efficiency messaging in the market place, socio-economic status, trust in the messaging, friends and families use of CFLs, product-related barriers, and political orientation. Figure 2 shows an example of our CFL model.

**Figure 2. Flex Your Power CFL Structural Equation Modeling Path Model**



By utilizing the SEM analytic approach, the Opinion Dynamics team estimated the relative strength of various causal paths through statistical tools and measures. The strength of this model lies in its ability to assess and remove measurement error from hypothesized structural relations. Through accounting for the presence of hypothesized variables in the models, the ODC team can examine a wider range of factors that cause an individual to take desired actions. This form of statistical analysis is perfectly suited for message testing, where a large number of factors may actually influence targeted behaviors.

Through SEM analysis, the ODC team can test the relative impact of program messaging on participants' energy efficiency actions. The SEM analysis assesses the programs' impact through three primary tests:

1. Determining a standardized coefficient associated with participants' exposure to the program messaging that totals direct and indirect effects of the messages after taking account of competing and correlated alternative influences. This total effects coefficient acts as a measurement of the relative impact of the messaging on the participants' behavior through accounting for competing variables (such as other messages in the marketplace, influence of friends and associates, etc). This figure allows the ODC team to develop an estimate of attribution from our survey efforts outlined below. Using this coefficient, our team can estimate the percentage of behavior change that can be attributed to the program messaging.
2. Testing the overall fit of the theoretical model to the empirical data. SEM employs a number of "goodness of fit" tests that indicates whether the theorized model accounts for the observed relations among variables.
3. Testing the "goodness of fit" of the SEM model when removing the influence of program messaging. Through this analytic, the ODC team can simulate the impact of the statewide program messages by setting them to zero in the model. The change in model fit when the program variable is set to zero will reflect the importance of program messages in influencing customer behavior.

By using the SEM approach, the ODC team can assess the program effects accounting for a wider range of potential influences on the behavioral outcome. In addition, the model can treat changes in attitudes and beliefs and intermediate outcomes, thus allowing our team to statistically assess the programs' ability to move participants along the behavior change continuum. Thus, the model allows us to control for various contributing factors to behavior change and assess the implications of each effect on the desired behavior change. These insights cannot be developed through standard effectiveness studies.

**Effectiveness Approach: Traditional Pre and Post Testing.** Pre and post testing of marketing and outreach program effects is the most commonly used form of effectiveness testing for advertising across all industries. Typically relying on RDD sampling methods, the use of pre and post testing is an important tool for quantifying gross programs effects over time. By using carefully planned and timed pre and post studies, evaluators are able to determine a baseline on a series of metrics central to a program’s success to compare to post studies which discern the sample’s movement after exposure to a program.

The Opinion Dynamics team used this traditional approach as the backbone of our evaluation effort, launching a comprehensive, statewide tracking survey aimed targeting California’s English and Spanish speaking populations. In addition, our team fielded survey efforts in two other states to serve as a comparison group for our California study. Each survey tested respondents’ attitudes, knowledge, program awareness, and behavior changes related to energy efficiency. By fielding this study repeatedly within and outside of the state of California, the ODC team is able to determine both changes over time within the California while adjusting for national and regional trends measured in our comparison groups.

This analysis, while commonly used, tells only part of the story--it doesn’t tease out the other effects that might lead to behavior change in California versus other states. Since every region that can be used for comparison purposes is not completely the same except for the advertising effort, we cannot assume that differences can be attributed specifically to marketing efforts and not to other factors. This is why our work is greatly enhanced when triangulating our findings with our verified reach assessment and our structural equation modeling effort.

**Qualitative Insights: Ethnographic Research.** In addition to the insights gained through both our efficacy and effectiveness studies, the ODC team conducted qualitative research to understand and better distinguish the potential effects of the Flex Your Power program. While a great deal of information may be gained through statistical modeling and innovative survey research, the ODC team understands that behaviors are understood as much through observations as they are through self report. Thus, our team drew on ethnographic research to supplement and add depth to our quantitative insights.

Ethnography is a field method of participant observation, developed by cultural anthropologists and widely used in the social sciences. The method aims to understand how meaning and behaviors are developed, maintained, and supported through micro-practices – that is, one’s everyday interactions with others. By applying this method, the ODC team can better to pay attention to what people *do*, not simply what people *say*.

To conduct this research, the ODC team attended a series of events throughout the state to evaluation the Flex Your Power program’s rural outreach efforts. These events aim to supplement the mass media campaign by using grass-roots organizing and events to target the hard-to-reach rural population. The program logic argues that such one-to-one interactions are the perfect channel to communicate about energy efficiency and inspire action. By attending these events, the ODC team was able to assess how such interpersonal communication facilitates behavior change by watching the physical reactions of the “participant” to the information that was provided. In addition, the team was able to observe how individuals exposed to this information “received” it while accompanied by other stakeholders in the decision-making process, such as a partner,

spouse, or child. These observations not only allow us to understand how participants are influenced by other stakeholders in their lives, but also how these interactions may or may not appeal to their needs based on how they appear to receive the information provided to them. When combined with traditional post-event intercept surveys, these observations can add depth to otherwise traditional research efforts.

**Innovative Method: Social Network Analysis.** In addition to our ethnographic research, the ODC evaluation team has conducted Social Network Analysis to inform marketing and outreach campaign development and planning. This method, as described by fellow panelist Caroline Wilson, may be used to understand networks of communication and information sharing to determine the best and most effective methods to disseminate information. In addition, this method can illuminate networks of like-minded individuals, allow evaluators to understand the social influences on behavior change.

**How this applies to other jurisdictions.** In some ways, California is unique in that the program evaluation budgets are large enough to explore a myriad of different research methods. In other jurisdictions, evaluators need to be more judicious in their evaluation research plans. One key consideration for all jurisdictions is to allow the evaluator to work with the marketing team up front to gather baseline information, concentrate on some core efficacy measures, and set up the plan for quantifying success. The marketing firm must have clear and measurable goals set forth before they launch a program effort. Equally important, key point is that marketing outreach efforts should include more than just direct energy savings as a goal.

Nationwide, more and more entities are spreading energy efficiency messages—whether it be Home Depot, Al Gore, EnergyStar or NBC News. Evaluating the effectiveness of any one campaign is difficult as people tend not to remember who sponsored the message. Structural Equation Modeling allows evaluators to tease out the affects of a marketing effort from other possible reasons why someone may have taken an energy efficient action—this methodology holds lots of promise in other jurisdictions.

By looking beyond standard pre/post effectiveness tracking studies and exploring new options for determining program success, other jurisdictions will be better positioned to capture the breadth of program successes and to quantify multiple effects, not just energy savings alone.

### **How do we measure success of Marketing and Outreach Programs?**

The paper aimed to answer the question: how do we measure the success of marketing and outreach programs? We strongly believe that a combination of approaches, including both efficacy and effectiveness approaches, is the best way to measure success. Most notably, we argue two primary points: (1) methods for evaluating marketing and outreach programs may be adapted from other social marketing and outreach evaluations such as those in public health; (3) marketing and outreach programs require a commitment to innovative measurement tools to better assess and quantify behavior change along a continuum, and should incorporate efficacy, effectiveness, and qualitative approaches to enhance how we come to measure success. Through combining carefully defined definitions of success and advanced methodological techniques, energy

efficiency evaluators will be better equipped to determine the relative value of marketing and outreach programs against one-another and standard resource acquisition programs. Thus, increasingly able to provide implementers and regulators with the feedback necessary to ensure that programs live up to their stated goals and expectations.

## **Bibliography**

Evans W.D., J. Uhrig, K. Davis, and L. McCormack. 2008. "Efficacy methods to evaluate health communication and marketing campaigns." *J Health Communications* 2:140-158

TecMarket Works Team. 2006. "California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals."

National Social Marketing Centre. 2007. "Big pocket guide to Social Marketing, Second Edition." <http://www.nsms.org.uk/public/default.aspx>. National Social Marketing Center. December 1<sup>st</sup>, 2008.

<[http://www.nsms.org.uk/images/CoreFiles/NSMC\\_Big\\_Pocket\\_Guide\\_Aug\\_2007.pdf](http://www.nsms.org.uk/images/CoreFiles/NSMC_Big_Pocket_Guide_Aug_2007.pdf).>

Wahlstrand, G., C. McRae, and A.C. Chapman. (In Review). 2008. "Studying pro-environmental behavior: The theory of planned behavior predicts intentions to use compact fluorescent light bulbs." *Environment and Behavior*.

Yang, Hongmei, and Peter Buyungo. 2007. "PERForMance Review: Behavioral Indicators." [www.psi.org](http://www.psi.org). Population Services International. December 1<sup>st</sup>, 2008.

<[http://www.psi.org/research/toolkits/performance\\_review.pdf](http://www.psi.org/research/toolkits/performance_review.pdf).>