

Evaluating Media Campaign Effectiveness: Others Do it, Why Don't We?

*Lori Megdal, Megdal & Associates
Sylvia Bender, California Energy Commission*

ABSTRACT

Advertising is often an important part of the strategy mix used to stimulate conservation actions and energy efficiency program participation. At the same time, advertising, along with broader energy-related social marketing campaigns, rarely receive the same level of evaluation effort as energy efficiency programs. Reliable estimates of advertising's impacts and effects on programs, useful information for campaign improvement, and data to assess optimal investment levels cannot be obtained without measurement and evaluation.

Social marketing techniques are widely used in influencing behavior in other social contexts, for example, food advertising (i.e., milk and egg campaigns), anti-smoking, anti-drug, disease control, health concerns, tourism, and the environment. This paper uses a "case study" approach to examine relevant media campaign evaluation methods used in these fields. The evaluation case studies provide quick overviews of the evaluation research designs and the usefulness of the findings.

Looking outside our own field provides the foundation for more effective evaluation of mass media campaigns in the energy efficiency field. Many of the research designs and evaluation methods could be applied in evaluating efficiency advertising.

Introduction

Advertising is often an important part of the strategy mix used to stimulate conservation actions and energy efficiency program participation. At the same time, advertising, along with broader energy-related social marketing campaigns, rarely receive the same level of evaluation effort as energy efficiency programs. Evaluators face several challenges in evaluating media campaigns: complexity, unpredictability of contextual events, confounding influences, lack of comparison groups, unfamiliarity with appropriate methods, and uncertainty about the explanatory value of outcomes. Most often, evaluators fall back on simply tracking frequency and reach. This lack of evaluative data, however, makes it difficult to support a case for costly advertising when facing funding decision trade-offs with more directly measurable programs. Without evaluations of energy-related advertising or social marketing efforts, any contributions they make to our overall efficiency efforts are unclear. It is time that the energy efficiency field conducted reliable effects evaluations of its advertising efforts. Reliable impact/effects evaluations of advertising could provide: 1) energy and demand savings estimates to be included in the portfolio savings; 2) estimates of the impacts on the probability of program participation; 3) more thorough information for improving campaign efforts; and 4) information for determining the optimal investment level.

Social marketing techniques are widely used in influencing behavior in other social contexts, for example, generic food advertising (i.e., milk and egg campaigns), anti-smoking, anti-drug, disease control, health concerns, and the environment. While little attention has been

paid to evaluating marketing campaigns in energy efficiency, other fields do offer constructive evaluation examples that could be useful for evaluating energy efficiency advertising. Selected well publicized signature evaluations from these other fields are presented as case studies herein.

Energy Efficiency Has Some Reputable Examples

The premise of this paper is that the energy efficiency field conducts few reliable effects evaluations on advertising efforts and could benefit by learning from other fields that have conducted good evaluations of advertising. Energy efficiency does have some well-known examples of reputable marketing effects evaluations to recognize, and perhaps a few others that have not been published and made widely available.

Advertising and educational programs dominated program efforts in the 1970s and early 1980s. Little evidence of the impact or effects of these advertising programs, however, is apparent thirty years later. If such studies do exist, it would be useful to our field to uncover them and to conduct a reliability assessment on their methodologies.

Wisconsin Focus on Energy conducted an effects evaluation for their “umbrella” advertising effort in 2001 and 2002. This is one of the few recent advertising evaluations in the energy efficiency field, and the only known example that proposed to examine responses over many time periods (a baseline random telephone survey and monthly tracking surveys reporting three-month moving averages). Significant funding reductions for the umbrella advertising campaign (for reasons outside of any evaluation results) led to a discontinuance of the planned effort. A more simple over-time comparison concerning a targeted CFL promotion, however, did occur. This more focused evaluation compared one region to the state overall, providing a quasi-experimental design that found evidence of the effectiveness of the advertising effort on awareness and reported CFL sales (Feldman & Rambo 2003).

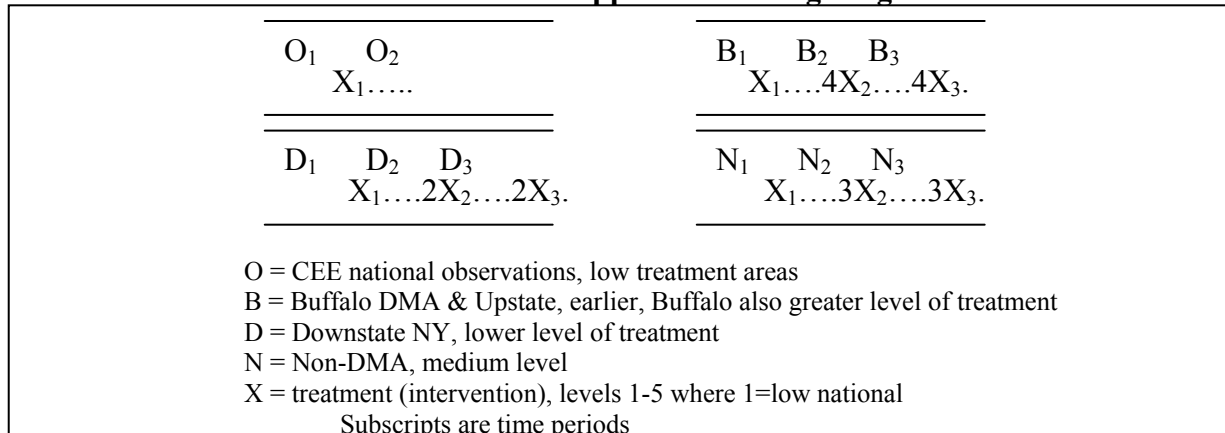
The Northwest Energy Efficiency Alliance constructed a mass marketing campaign to increase demand for efficient commercial buildings in the late 1990’s. An effects evaluation documented program-induced effects, but also recommended a significant change toward the use of more targeted messages rather than a mass marketing campaign. The program adopted these changes. The evaluation that made these findings and recommendations is summarized in an article from the 2001 IEPEC conference (Gordon, Peters & Dethman 2001).

There are two articles of note from earlier evaluations conducted for the New York State Energy Research and Development Authority (NYSERDA). A causality assessment in 2001 made comparisons over just two time periods, but investigated varying the intensity of advertising effort in four different geographic areas: (1) low national effort within the Consortium for Energy Efficiency (CEE) survey, (2) the lowest mass media effort for their program in downstate New York, (3) the effort experienced in New York’s non-designated market areas (non-DMAs), and (4) an earlier and more intense effort employed in the Buffalo, New York area. The research design for this evaluation is shown in Figure 1. The patterns of effects all differed by the advertising intensity level just as hypothesized, providing evidence of advertising as causality for the impacts (Megdal et al. 2001).

The second NYSERDA work referenced here may be one of the first to quantify energy demand (kW) impacts. The initial evaluation goal focused on testing the possibility that impacts from the \$2.5 million advertising effort could be quantified. A pre/post telephone survey on behavior surrounded a short, intense summer advertising campaign. Linking the survey data together with energy usage information and engineering algorithms produced estimates of load-

shifting attributable to the program. The evaluation estimated that the advertising campaign resulted in approximately 94 MWs of load shifting away from peak hours, a significant contribution to their portfolio (Engel et al. 2003).

Figure 1. Quasi-Experimental Design within the 2001 Causality Assessment of the New York ENERGY STAR® Appliance and Lighting Markets



Source: Megdal et al. 2001. Research design diagram based upon classic graduate textbook for research design: Cook, Thomas D. and Donald T. Campbell (1979) *Quasi-Experimentation: Design & Analysis Issues for Field Settings*, Houghton Mifflin Company, Boston, MA.

In 2003-2004, California sponsored the creation of a National Energy Efficiency Best Practices Study. This report conducted significant research to develop best practices for each of 15 program areas, including advertising and mass marketing. The four best practices identified for program evaluation of efficiency advertising efforts are:

1. Use an evaluation team that includes both efficiency program evaluators and experienced advertising specialists
2. Use standard advertising evaluation methods combined with periodic effects evaluations
3. Periodically conduct process evaluations
4. Involve program staff in the evaluation process and create a culture whereby evaluation findings are valued and integrated into program management

The executive summary includes a list of seven efficiency advertising program pitfalls. One of these pitfalls is particularly relevant for the issue presented in this paper: “Failure to understand the value of advertising effectiveness evaluation methods” (Quantum Consulting & Megdal 2004, O1.4-O1.5).

Advertising Evaluation “Case Studies”

A variety of significant and well-studied social advertising campaigns in fields other than energy are examined using a case study approach. Each of these case studies is documented by published papers and/or reports and only the briefest summary could be included here. The evaluation case studies are designed to summarize information that could be used in evaluating efficiency marketing campaigns. The questions posed for our “case studies” are as follows:

- What were the primary goals of the advertising campaign(s)?
- What was the evaluation research design? How were the net impacts of the advertising estimated?
- How did key findings provide evidence of the evaluation's effectiveness?

Case 1: Generic Advertising of Food Products

Generic advertising of food products has a large body of evaluation research. These are campaigns, often funded through mandatory assessments of producers, which promote the food product without reference to any specific brand, manufacturer or producer. Some of the most recognized of these include the “Got Milk” campaign, the “Incredible Edible Egg” and “Pork, the other white meat.” In fact, the recognizability of the “Got Milk” campaign has led to numerous copy-cat campaigns.

The critical effectiveness factor in most of these campaigns is a positive cost-benefit ratio. A recent economic analysis of milk advertising in eastern Canada provides an appendix citing 31 evaluative studies of generic advertising for a wide variety of food commodities that demonstrate prevailing benefit-cost ratios over 1 and many from 5 to 30 (Kaiser, Cranfield & Doyon 2006, Appendix V).

Research Design

Many of these studies have been conducted by the National Institute for Commodity Promotion Research & Evaluation (NICPRE) at the Department of Agricultural, Resource, and Managerial Economics at Cornell University or graduates of this program. Much of the current work follows the econometric supply and demand analysis models with disaggregated products and separate wholesale and retail supply/demand components as developed by Harry Kaiser in his 1995 study on generic dairy promotion in the U.S. (Kaiser 1995). In this study, he cites a summary study from 1991 which examined results from 47 studies that estimated the effects of generic milk advertising. In general, these econometric regression models predict quantity, demand and price based upon supply shifters at the wholesale and retail level to include price, advertising and the price of a substitute good (e.g. juice, soda), sometimes with numerous regression equations and various other variables depending upon the analysis being conducted (e.g., during price supports the government purchase price was included, Kaiser 1995).

The Kaiser, Cranfield and Doyon 2006 report used quarterly data from 1990 through 2004 to estimate the following demand model for milk to evaluate advertising impacts in each Canadian province:

$$\ln \text{SALES}_t = \alpha_0 + \alpha_1 \ln (\text{PRICE}_t) + \alpha_2 \ln \text{INCOME}_t + \alpha_3 \ln \text{SUBPRICE}_t + \alpha_4 \ln \text{MILKAD}_t + \alpha_5 \ln \text{NON-MILKAD}_t + \alpha_6 \ln \text{DUMQ}_t$$

where the SALES variable is the quarterly per capita fluid milk sales, PRICE is quarterly retail milk price index, SUBPRICE is the substitute price index for substitute goods, INCOME is the quarterly per capita disposable income, MILKAD is primary variable of interest – the generic milk advertising expenditures, NON-MILKAD is expenditures on generic promotion (e.g., brochures, nutritional education efforts, and other non-advertising promotions and public relations), DUMQ are quarterly dummy variables. Monetary variables were generally deflated by the consumer price index (CPI), except the advertising expenditures which were deflated by a separate advertising inflation index (given a significantly higher advertising inflation index occurring than that for the CPI).

Results

As stated above, most of the commodity promotion evaluation studies found positive benefit-cost ratios for the investments. Generic milk advertising had positive and statistically significant results on milk demand in all three regions and non-advertising marketing elements as a group had positive effects on demand in two of the regions. The average producer rate of return for the generic milk advertising was calculated to be \$2.2 for every dollar invested in the Maritime provinces, \$3.4 in Ontario and \$7.5 in Quebec. Analysis also addressed optimal levels of investment in advertising versus the non-advertising marketing elements (Kaiser, Cranfield & Doyon 2006).

Case 2: The U.S. Youth Anti-Smoking “Truth” Campaign

The American Legacy Foundation initiated youth anti-smoking campaigns in 2000 using funding obtained from the Master Settlement Agreement between tobacco companies and 46 states. Expenditures through 2002 on the national “truth” campaign averaged \$100 million per year following a very successful Florida “truth” campaign.

Prior campaigns using overt messages to tell youth not to smoke, or citing health or cost consequences to direct them not to smoke, had not proven effective. The “truth” campaign, however, was designed around the fact that teenagers naturally desire to rebel as part of their differentiation and growing up process. The “truth” campaign wraps the health consequences of smoking around the fact that the tobacco industry purposely tries to manipulate people to become smokers. The campaign utilizes the natural teenage rebellious phase to allow them to rebel against this manipulation.

Impact Evaluation Research Design

The study used logic models, process evaluation, and outcome evaluation through a pre/post quasi-experimental design. Much of the earlier work used an annual national telephone survey of young people ages 9 to 13 (the Youth Media Campaign Longitudinal Survey-YMCLS) supported by the Centers for Disease Control and Prevention (CDC). The two primary data inputs for the 2005 Farrelly et al. evaluation summarized here were a national spring survey of youth (Monitoring the Future, MTF) and a measurement for the level of exposure. This survey was conducted by the University of Michigan’s Institute for Social Research and included 18,000, 17,000 and 16,000, 8th, 10th and 12th grade students, respectively from approximately 420 public and private secondary schools. The survey used a multi-stage random sampling design to ensure proper national representation via geography, school and classes within the schools.

Though the campaign involved a national media buy, variation in exposure occurred due to the presence or absence of local affiliates and cable markets for the networks airing the advertisements (FOX, UPN and WB). Cumulative gross rating points (GRPs) for each of the 210 television markets in the United States provided measurement of student exposure to the campaign. (GRP is reach times frequency, which includes the percentage of the audience exposed and the number of times they are exposed to the advertising.) The variation over the two-year period evaluated was from 3,867 GRPs to 20,367 GRPs and the analysis included pre-exposure survey results where the GRPs were zero (Farrelly et al. 2005). This technique of

measuring effects by comparing those with high exposure to the campaign to those with low exposure is known as dose-response design.

Multivariable logistic regression was used to analyze the GRPs on current youth smoking prevalence, controlling for individual, media market and state level confounders. Regressions were run by both individual and combined grade levels. GRPs were scaled so odds ratios could be developed per 10,000 GRPs. (Most health and medicine evaluations report results as odds ratios. It is the ratio of the odds of an event occurring in one group to the odds of it occurring in another group, which is often the control group.)

Results

The impact evaluation determined that the campaign accounted for a significant portion of the decline in youth smoking prevalence during this time period. Between 1999 and 2002, smoking prevalence among students declined from 25.3% to 18.0% and the campaign accounted for 22% of this decline. A statistically significant dose-response relationship was found between “truth” campaign exposure and current youth smoking prevalence with an odds ratio of 0.78. Separate regressions by grade showed the largest effects for 8th-grade students with an odds ratio of 0.61 ($p < .05$).

Case 3: The National Youth Anti-Drug Ad Campaign

The National Youth Anti-Drug Media Campaign (Campaign) launched in 1998 with a 5-year federal appropriation of nearly \$1 billion and \$50 million for evaluation. Structured as an integrated social marketing and public health communication campaign, the Campaign focused primarily on preventing and reducing drug use among youth and encouraging parents to play a more active role in the process. The Campaign reached its intended audiences through paid commercial advertising, grassroots public outreach and specialized supporting communication efforts. After an initial pilot intervention effort in 12 cities, the Campaign evolved in complexity and intensity to a fully implemented nationwide multiple media campaign which ran from 1999 through mid-2003. Semiannual evaluation reports conducted by Westat and the Annenberg School of Communication provided increasingly robust detail on estimates of association and change culminating in an examination of the longitudinal and delayed-effects of Campaign exposure on cognitive and behavioral outcomes (National Youth Anti-Drug Media Campaign 2000-2005).

Two evaluation methods used in this campaign are particularly noteworthy: theory-based evaluation and propensity scoring. The Campaign illustrates the use of a theory-based evaluation effort that incorporates known evidence about what influences drug use and basic theory about communication and health behavior change into an evaluation design consistent with the stated theory of effect. Theory-based evaluations are especially useful for investigations into dynamic and multi-faceted processes, such as behavior change. Secondly, the “propensity scoring method” reduces the risk that observed differences over time are the result of confounding variables.

Theory-Based Evaluation as a Research Design

In order to gauge the impact of the Campaign on awareness, knowledge, attitudes and behavior, the evaluation asked: (1) Is the Campaign reaching its target audience? (2) Are the desired changes going in the right direction? (3) How much of the observed changes can be attributed to the Campaign? (4) What is learned to support ongoing decision making for the Campaign?

A series of influence diagrams modeled the routes through which the Campaign might influence actions of both youth and their parents. For example, the anti-drug messages might be received through multiple channels: exposure to the message, interaction with friends, interaction with parents and interaction with organizations. The Campaign evaluation model also mapped a series of assumptions about how exposure to anti-drug messages might be turned into behavior using the theory of reasoned action (Fishbein & Ajaen 1975), supplemented by the idea of self-efficacy (Bandura 1986). The model further incorporated a wide range of external (i.e., parental monitoring, friends' behaviors) and personal factors (i.e., sensation seeking, ambitions) that have been shown to influence drug use. The final elements of the theory incorporated the possibility of time lags between campaign activities and effects, and possible generalization from specific campaign messages to other behaviors.

The National Survey of Parents and Youth, conducted in seven discrete waves from September 1999 through June 2003 served as the primary data collection tool. The survey collected initial and two rounds of follow-up data from national representative samples of youth between 9 and 18 years of age (8,000 individuals) and their parents (5,500 individuals) in their homes using touch-screen laptop computers to mitigate privacy concerns associated with the sensitive nature of the questions. Evaluators measured both general exposure to anti-drug information and specific exposure to campaign messages.

The influence models mapped how a successful campaign would bring about the desired changes, which in turn determined the range of questions for the survey questionnaires. The specific measurement objectives focused on changes in the drug-related knowledge, attitudes, beliefs and behavior of youth and their parents, relationships between those effects and media exposure, association of change in parents with change in youth, and comparisons of high exposure groups to low exposure groups. The research design notes that: "The way that media campaigns can affect behavior is often complex, and if that complexity is not reflected in the evaluation design, many of the effects may go undetected (Hornik & Yanovitzky 2003, 222)."

Measuring Attribution Using Propensity Scoring Methodology

Making any estimates of the Campaign's effects on audiences' cognitive and behavioral outcomes, however, is risky without controlling for confounding variables. Confounders are variables that can explain variations both in the level of campaign exposure and in outcomes, but that are not themselves caused by the campaign exposure or outcomes (for example, previous drug use, family characteristics and personal traits). Propensity scoring methodology, developed by Rosenbaum and Rubin in 1983 and used widely in applied medicine and biostatistics (e.g., Rubin 1997), and psychology (e.g., see review of method and issues in Luellen, Shadish & Clark 2005), is a sophisticated statistical technique useful for improving causal inferences from observational data by adjusting any effects for possible selection bias. The most critical step in using this technique is to base selection of the confounding variables on theoretical grounds and

previous empirical evidence of relationships *a priori* of analysis of the evaluation data. These variables are those that both increase the probability or degree of campaign exposure and affect the desired outcome. It is important to distinguish these from mediator variables, variables that link the campaign exposure to the outcomes, but do not increase the probability or degree of exposure. The selected confounders are tested for validation (often using two-way analysis of variance (ANOVA)). The propensity score is then estimated for each subject. This has been done using logistic regression, discriminant analysis, probit models, and classification trees (often using bootstrapping methods), the most common of which is logistic regression. The dependent variable is the probability of being exposed to the campaign based upon the confounder variables. Matched pairs or strata that balance the covariates are constructed and the treatment effects are analyzed for these overlapping strata.

Results

Recall of the television advertising climbed over the 3.5 years of the fully implemented campaign, reaching 51 percent among parents and 52 among youth. Both groups reported substantial recognition of the Campaign's "anti-drug" brand phrases. These early signs of success, however, were limited to increased awareness.

Evidence consistent with a favorable campaign effect on parents did continue throughout the Campaign, specifically for talking with children about drugs, doing fun activities with children, and beliefs about monitoring children. Evidence about changes in parents' monitoring behavior, however, proved much weaker. No evidence was found for favorable indirect effects on youth behavior or beliefs as the result of parents' exposure to the Campaign.

Little evidence of direct favorable Campaign effects on youth was found. For the primary audience of 14- to 16-year olds, past-year drug use increased from 2000 through 2003. Youth who were more exposed to the Campaign messages were no more likely to hold favorable beliefs or intentions about marijuana than youth less exposed to those messages over the entire course of the Campaign.

Case 4: Return on Investment (ROI) Evaluations of Ads for Tourism

Research Design

The sophistication of the evaluations and studies for economic impact of advertising on tourism has increased over the last five years. Often these include pre-wave and post-wave surveys with the targeted audiences. Traditional inquiries include awareness, travel intentions and attitudes associated with the campaign. Tourism campaigns often seek to generate inquiries and disbursement of informational packets. Surveys of those making those inquiries are conducted several months afterwards to assess travel made, booked, or planned compared to prior travel plans to the targeted destination. The survey results are used to create a conversion rate, the proportion of travelers visiting the target location that would not otherwise have occurred. The surveys also ask travel modes, stay and expenditures by type. These are then used to calculate expenditures and the return on investment (ROI) for the advertising investments.

Results

A 2003 study sponsored by the Canadian Tourism Commission used a combination of these methods: a post-wave only survey, and an on-line panel survey. The latter study obtained estimates from television advertising of additional revenues due to unplanned or switched trips to Canada of \$CN 36 million and another \$CN 56.9 million due to being influenced to take a longer trip.

An evaluation firm specializing in evaluating travel and tourism efforts, Longwoods International, developed and markets the Accountability Method Longwoods R.O.EYE™. Using surveys and an econometric analysis method, they conduct impact evaluations of tourism advertising campaigns which produce return on investment (ROI) statistics. Their work has been recognized as a best practice in a study funded by the U.S. Department of Commerce. In a study on the Michigan 2004 tourism advertising campaign, Longwoods International estimated trips generated per media dollar (0.28), travel spending per media dollar spent (\$46.51) and state taxes generated per media dollar spent at \$3.27 (which would translate to a simplified benefit/cost ratio of over 3).

Other Media Campaigns Using Impact/Effects Evaluations

Noteworthy impact/effect evaluations exist in many other areas of advertising. It is beyond the length and scope of this paper to provide even summaries of these, but it is worth recognizing the broad range of advertising efforts receiving significant evaluation.

The U.S. military conducts impact evaluations on the effectiveness of its recruitment campaigns. Two leading research organizations recently conducted an evaluation of joint versus service-specific recruitment advertising. An experimental design varied the levels of joint and service-specific advertising across the country and observed the number of recruits obtained. The initial use of statistical-econometric analyses could not definitively determine which form of advertising was more effective. The Operations Research Center of Excellence (ORCEN) using the same data but adding information on managerial efficiency in the recruiting districts, conducted a revised evaluation U.S. Military Academy's Operations Research Center of Excellence. Specific impact results, including and controlling for differences in managerial efficiency, resulted in a finding of much greater effectiveness from the joint service advertising.

An ongoing public health evaluation, sponsored through the Centers for Disease Control and Prevention (CDC), focuses on the VERB™ youth media campaign to get 9-13 year olds to engage in physical activity. Evaluation activities include formative research, a program logic model, process evaluation, and outcome evaluation using the Youth Media Campaign Longitudinal Survey (YMCLS).

Researchers from Claremont Graduate School have recently completed impact evaluations of Hispanic mass media campaigns to promote organ donation (Alvaro et al., 2005). The research design highlights the importance of simultaneously working with program message development and testing.

Conclusions and Recommendations

There are many impact evaluations of advertising efforts in other fields that obtain meaningful and statistically significant results. The span and depth of these suggests that there is

no reason why meaningful and significant impact and effect evaluations can not be conducted on energy efficiency advertising efforts.

Many different research designs and evaluation methods are used across the advertising evaluations reviewed here. The review and references provide examples that could prove to be useful approaches for evaluating efficiency advertising.

One of the themes seen over a number of the studies is the importance of understanding the expected effects from the advertising and the linkages through various mechanisms (mitigating variables) and alternative drivers (confounding variables) to longer-term effects. Logic models help clarify linkages among program components, focus on short-term, intermediate, and long-term outcomes and guide appropriate data collection and analyses. This has been recognized as important in the anti-drug campaign evaluation, the military recruitment evaluation, and as seen in the program logic efforts for the youth anti-smoking effort, and the exercise and activity campaign VERBTM. There is also evidence in these works of the importance of understanding the underlying behavioral models assumed and how the advertising logic (program, messages and delivery) works at this greater level of detail.

Advertising evaluations have made use of differential treatment across areas and time. Research designs with multiple comparisons provide a greater ability to measure effects and the level of effect generated per increment of advertising than a simple pre-post survey design. Experimental design could be incorporated into large-scale media plans to accomplish this with even greater reliability. Alternatively, dose-response type of research design could be used more often with detailed examinations of how the advertising efforts were actually rolled out.

Skeptics often charge that because advertising changes demand, it also changes prices, thereby making it impossible to evaluate its effectiveness. The supply-demand modeling used to evaluate generic food advertising, however, provides examples of methods specifically designed to measure the effect of advertising on prices and the overall dynamic supply-demand system. The numerous studies of commodity advertising provide examples and discussions of issues that must be addressed when using this type of evaluation method.

These case studies document a variety of sophisticated statistical methods being used to evaluate advertising, which range from supply-demand modeling in generic food advertising, to multi-variant logistic regression in the youth anti-smoking campaign, and to propensity scoring analysis as a control for selection biases in the youth anti-drug campaign. Lessons on evaluation methods from these fields appear to be fruitful additions to the toolbox of energy efficiency evaluators.

References

- Alvaro, Eusebio M., Jason T. Siegel, William D. Crano, and Alexander Dominick. 2005. "Evaluating Mass Media/Community Outreach Campaigns: Challenges and Outcomes Associates With a Campaign Targeting Hispanic Organ Donation" and "Increasing Organ Donation: Evaluating Potential Mass Media Appeals Before Spending Mass Dollars". Presentations at the 2005 Canadian Evaluation Society (CES) and American Evaluation Association (AEA) Joint Conference in Toronto, Canada.
- Bandura, Albert. 1986. *The Social Foundation of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, New Jersey: Prentice-Hall.

- Booze Allen Hamilton and University of Pennsylvania's Wharton School. 2003. "When Art Meets Science: The Challenge of ROI Marketing", *Strategy+Business*: December 17.
- Canadian Tourism Commission. 2003. *Advertising Evaluation: Approaches Adopted & Lessons Learned – The Experience of the Canadian Tourism Commission*.
- Coffman, Julia 2002. *Public Communication Campaign Evaluation: An Environmental Scan of Challenges, Criticisms, Practice and Opportunities*. Prepared for the Communications Consortium Media Center, Harvard Family Research Project. Cambridge, Mass.: Harvard University: May.
- Engel, Victoria S, Lori Megdal, Tom Rooney, Lawrence Pakenas, and Susan Soweck. 2003. "Quantifying Load-Shifting Benefits from an Advertising Campaign." In *Proceedings of the 2002 International Energy Program Evaluation*: 393-401. Madison, Wisc: International Energy Program Evaluation Conference.
- Farrelly, Matthew C., Kevin C. Davis, M. Lyndon Haviland, Peter Messeri, and Cheryl G. Heulton. 2005. "Evidence of a Dose-Response Relationship Between "truth" Anti-smoking Ads and Youth Smoking Prevalence." *American Journal of Public Health* 95 (3): 425-31.
- Feldman, Shel and Eric Rambo. 2003. "How'm I Doing? Tracking the Effectiveness of Advertising an Energy Efficiency Program." In *Proceedings from the 2003 International Program Evaluation Conference*: 403-16. Madison, Wisc: International Energy Program Evaluation Conference.
- Ferrero, Jennifer, Leen Boon, Harry M. Kaiser, and Olan D. Forker. 1996. *Annotated Bibliography of Generic Commodity Promotion Research (revised)*. NICPRE 96-03. Ithaca, New York: National Institute for Commodity Promotion & Research Evaluation (NICPRE), Department of Applied Economics and Management of Cornell University.
- Fishbein, Martin and Icek Ajzen. 1975. *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, Mass.: Addison-Wesley.
- Gordon, Jane, Jane S. Peters, and Linda Dethman. 2001. "Evaluating a Campaign to Increase Demand for Energy Efficient Commercial Buildings." In *Proceedings of the 2001 International Energy Program Evaluation Conference*: 695-705. Madison, Wisc.: International Energy Program Evaluation Conference.
- Hornik, Robert, David Maklan, Diane Cadell, Carlin Henry Barmada, Lela Jacobsohn, Amalia Prado, Anca Romantan, Robert Orwin, Sanjeev Sridharan, Elaine Zanutto, Robert Baskin, Adam Chu, Carol Morin, Kristie Taylor, and Diane Steele. 2002. *Evaluation of the National Youth Anti-Drug Media Campaign: Fifth Semi-Annual Report of Findings*, prepared for the National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services. Rockville, Maryland: Westat and Annenberg School of Communication of the University of Pennsylvania.
- Hornik, Robert and Itzhak Yanovitzky. 2003. "Using Theory to Design Evaluations of Communication Campaigns: The Case of the National Youth Anti-Drug Media Campaign," *Communication Theory*, May: 204-24.

- Kaiser, Harry M. 1995. *An Analysis of Generic Dairy Promotion in the United States*. NICPRE 95-02. Ithaca, New York: National Institute for Commodity Promotion & Research Evaluation (NICPRE), Department of Applied Economics and Management of Cornell University.
- Kaiser, Harry M., John Cranfield, and Maurice Doyon. 2006. *An Economic Analysis of Generic Fluid Milk Advertising in Ontario, Quebec, the Maritime Provinces*. NICPRE 06-01, R.B. 2006-01. Ithaca, New York: National Institute for Commodity Promotion & Research Evaluation (NICPRE), Cornell University.
- Longwoods International. 2005. *Michigan Image and 2004 Advertising Evaluation*. Toronto, Canada.
- Luellen, Jason K., William R. Shadish and M.H. Clark. 2005. "Propensity Scores: An Introduction and Experimental Test," *Evaluation Review*: 29 (6): 530-58.
- Megdal, Lori, Mark Coleman, Jennifer Ellefsen, Larry Pakenas, Helen Kim, and Scott Albert. 2001. "He did it! He did it! – Providing Evidence for Causality." In *Proceedings from the 12th National Energy Services Conference*: 140-50. Association of Energy Services Professionals.
- National Youth Anti-Drug Media Campaign. 2000-2005. Evaluations can be found online at <http://www.mediacampaign.org/publications/performance.html> Washington, D.C.: Office of National Drug Control Policy.
- Quantum Consulting and Lori Megdal. 2004. *National Energy Efficiency Best Practices Study, Volume 01- Cross-Cutting, Advertising Best Practices Report*, prepared for the California Best Practices Project Advisory Group, submitted to Dr. Kenneth James of Pacific Gas & Electric, December. Report available on www.eebestpractices.com. This chapter can be found under Program Areas – Cross-cutting—Mass Market Advertising.
- Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70 (1): 41-55.
- Rubin, Donald. 1997. "Estimating Causal Effects from Large Data Sets Using Propensity Scores," *Annals of Internal Medicine* 127 (8): 757-63.
- Schmit, Todd M., J. Carlos Reberte and Harry M. Kaiser. 1998. "An economic analysis of generic egg advertising in California, 1985-1995," *Agribusiness* 13 (4): 365-73.
- U.S. Military Academy's Operations Research Center of Excellence (ORCEN). *An Evaluation of Joint and Service Specific Advertising Efficiency for Military Recruitment* for the United States Army Recruiting Command (USAREC). Available online: http://www.orcen.usma.edu/Research%20Projects/Previous%20Projects/AY01/recruitment_evaluation.htm
- Yanovitzky, Itzhak, Elaine Zanutto, and Robert Hornik. 2005. "Estimating causal effects of public health education campaigns using propensity score methodology," *Evaluation and Program Planning* 28: 209-20.