How Much Is That Training Program Worth?
Quantifying the Value of Training and Other 'Fuzzy' Education Events

Ingo Bensch, Scott Pigg, and Marge Anderson, Energy Center of Wisconsin

ABSTRACT

The bottom line . . . it drives business decisions on everything from the peanuts offered on short flights to the design of new buildings. Energy-related decisions are often also evaluated based on costs and benefits, which is why the energy efficiency industry has learned to speak the language of payback and non-energy benefits. Those energy measures (or anything else) whose value can be quantified stand a better chance of being funded and implemented.

This reality is driving the training and meetings industry toward measuring the value of their offerings and stating them in business terms. This industry is beginning to adopt return on investment (ROI) calculations as a standard by which businesses can measure such formerly fuzzy necessities as training, conferences, and meetings. Several industry providers are applying a rigorous, empirically-based methodology to such events in an effort to quantify their value.

This paper describes a self-funded, internal assessment of return on investment for a Midwestern energy efficiency conference. Through a combination of interviews, follow-up surveys, program tracking, and content analysis, we attempted to monetize the value of the conference to show whether the sponsorship investment was "worth it" in terms of quantifiable energy impacts.

In this paper, we explain our approach, share our results, highlight lessons learned, and argue that our modest effort suggests that education and training impact studies warrant greater attention by the energy efficiency industry.

Introduction

In many jurisdictions, energy efficiency funding flows to the programs and activities that can demonstrate energy savings. For example, California's energy efficiency programs are held up to a standard that emphasizes measurable resource acquisition (Eley et al. 2004). The New York State Energy Research and Development Authority's programs have been increasingly emphasizing evaluated energy savings – especially peak reductions – and Wisconsin's Focus on Energy program is held to a standard that requires a 1:1 benefit/cost ratio using only traditional measures of energy efficiency (York et al. 2002). This funding reality has the unintended side-effect of devaluing programmatic activities whose impacts are indirect or not measured.

Education and training of market actors whose practices affect energy usage is an important example of a programmatic activity whose impact generally is not measured adequately (Green and Skumatz 2000). Education and training are included in program plans even without measured energy savings because they are one way of addressing informational barriers, but they are also easily scaled back or cut in favor of those activities that bring in the ever-important verified kilowatt-hours and therms.

Barriers to measuring the impact of training are numerous. They include:

- the absence of an expectation that training provides measurable energy savings, which causes program managers and evaluators not to look for savings;
- the perception that training impact evaluations are expensive;
- a lack of established practices within the energy industry for measuring training impacts;
- the challenge of attribution when training is usually provided in conjunction with other programmatic activities; and
- a lack of familiarity among program managers and evaluators with adult learning methodology and training design.

However, the impact of training has been measured in other industries – and even within energy efficiency – and shown positive results. For over 20 years, the training and development industry has contributed a rich body of work around training impact measurement, evaluation, and return on investment in both the public and private sector (Phillips 2003) (Kirkpatrick 1994). More recently, the meetings industry has made efforts to study the impacts of the events that participants in that industry organize.[1]

Perhaps the two most compelling examples of this measurement in the energy efficiency industry are the body of evaluation work around the Building Operators Certification (BOC) training and the Department of Energy-funded evaluation of the Compressed Air Challenge (CAC) training. The most recent results for BOC training included savings estimates of 0.35 kilowatt-hours and 0.34 million BTU per participant per square foot (RLW Analytics 2005). The CAC training saved 149 megawatt-hours per participant, an average savings per attendee implementing measures of $7,428, and a sponsor benefit/cost ratio of 82:1 (Lawrence Berkeley National Laboratory and Xenergy, Inc. 2004).

This paper describes an impact study of a technical training conference for which energy-based impacts were highly positive and discusses its methodology, its results, and its application to other energy efficiency training activities.

Return on Investment Study for the Better Buildings: Better Business Conference

The Energy Center of Wisconsin has produced a technical training conference for cold-climate residential building professionals annually since 1999. It is currently known as the Better Buildings: Better Business Conference, and its primary objective is to increase the energy efficiency of Wisconsin's homes. The conference uses many methods to achieve this objective:

- skills-based, hands-on workshops on a variety of best practices in energy efficient home construction, including shell and foundation insulation, framing, and mechanical ventilation;
- focused training in building science concepts illustrated with cold-climate examples by regional and national experts;

[1] See www.astd.org and www.mpiweb.org for more information.

- prominent integration of the Wisconsin ENERGY STAR™ programs[2] into the curriculum, general session, and trade show as well as satellite events like a new construction awards banquet;
- use of instructors who have been specifically trained by the Energy Center in training design for behavior change;
- a trade show with hands-on technology demonstrations to help attendees understand the functionality of energy-efficient technologies; and
- celebration of a community of practice around energy efficiency – providing recognition through two awards programs, training showcases, and networking – for builders and subcontractors practicing leadership in energy efficiency.

Given the conference's design intent, in 2005 the Energy Center decided to study the financial return to the conference from the perspective of two key stakeholders: Wisconsin's ENERGY STAR programs and the Energy Center itself. We sought to address two key research questions. First, did the conference have an impact on job-site practices or product selection? Second, were the recognition and community-building activities creating more motivation to maintain builders' commitment to energy efficiency?

As a self-funded effort, we did not have the resources for a comprehensive assessment of the conference. Also, our goal was simply to get a rough gauge on the extent to which we could ascribe behavior change to the conference to see whether additional efforts to isolate conference impacts might be warranted.

Implementing the Study

We implemented our study in four steps:

1. We developed hypotheses of various ways the conference might bring quantifiable value to the sponsors being studied;
2. We contacted a sample of businesses that attended the conference to identify self-reported changes in practices and purchases from the conference;
3. We attempted to quantify the energy savings (or other return to the sponsor) from changes identified in Step 2; and
4. We extrapolated the savings we identified from our sample over time and across all participants to approximate the total impact from the conference.

Hypotheses

Because we were assessing the return of the conference to two of its sponsors, we began by examining the organizational goals of those sponsors and developing hypotheses of how the conference might contribute to these goals. These hypotheses included a behavior change model leading to energy savings, marketing-oriented benefits for a sponsoring program, and even intangible benefits such as increased recognition for, and enthusiasm by, program staff. We divided these into hypotheses we could measure and those we could not. The most important quantifiable hypotheses were:

[2] Specifically, the Wisconsin ENERGY STAR Homes and Home Performance with ENERGY STAR programs.

- Building professionals learn new ideas or practices at the conference that they implement into subsequent homes, thereby changing the energy consumption of the homes' future occupants.
- Extensive visibility for the Wisconsin ENERGY STAR programs causes more builders and remodelers to join the programs (or to join them sooner) and modify their building practices to meet program requirements, thereby changing the energy consumption by their homes' occupants.

Other hypotheses – which we labeled as unquantifiable – included the following:

- The conference's sessions on business practices causes builders and others to become more successful on the business side of their operation, which results in greater numbers of homes run through the sponsoring programs, thereby resulting in some incremental energy savings.
- The conference motivates and prepares more people to be consultants for the sponsoring programs and increases the number of consultants "selling" the programs, and thereby causes more homes to be run through the programs, which results in some incremental energy savings.
- Growth in sponsoring programs' brand recognition and increased participation gives the programs more market influence, which contributes to program sustainability and future growth.
- The same chains of events that can lead to energy savings also impact non-energy goals of the sponsoring programs, like safety, durability, and comfort.

These hypotheses are not necessarily unquantifiable, but the lack of established and accepted practices for quantifying some of these outcomes and challenges in attribution would make it more difficult to produce defensible numeric results for these hypotheses than those we chose to quantify.

Assessing Behavior Change

The most important hypotheses were based on some sort of behavior change by conference attendees, so our next step was to estimate that behavior change. We did so by designing and implementing structured interviews with a sample of conference attendees in order to elicit information about how the conference might have changed the field practices of attendees.

The sample was drawn from conference registration data. Table 1 shows our estimate of the number of various types of firms represented at the training, as well as the number of each type included in our sample and the number of interviews completed. We sampled randomly from the firms represented at the training that met the following criteria:

- Wisconsin-based organization;
- identifiable type of business (based on conference staff knowledge of the firm or clear identification of the business type from the company's name);
- telephone number in the conference registration database; and

- either a builder, remodeler, subcontractor (i.e. HVAC contractor, insulator, etc.), or consultant to the Wisconsin ENERGY STAR programs.

Table 1. Conference Attendance and Sampling

Type of Business	Estimated Number in Attendance	Number Sampled	Number Interviewed
builders	80	24	16
remodelers	12	2	4[3]
subcontractors	40	11	4
consultants	32	13	7
other	7	0	3[4]
uncertain type	31	excluded	0
no telephone number	8	excluded	0
out-of-state	38	excluded	0
secondary audience	97	excluded	0

In all, we completed structured telephone interviews with 34 out of 50 residential building firms we sampled from the list of all firms represented at the conference. These survey completions represented 20 percent of all builders who attended, 33 percent of all remodelers, and 10 percent of all subcontractors. Results presented here are based primarily on self-reports from 24 of the 34 attending businesses we interviewed (the builders, remodelers, and subcontractors among the interviewees), but our analysis was supplemented with helpful background information from day-of-event evaluations, an analysis of conference session content, and review of attendee responses for reasonableness. Figure 1 shows our data sources (on the left side) in graphical form.

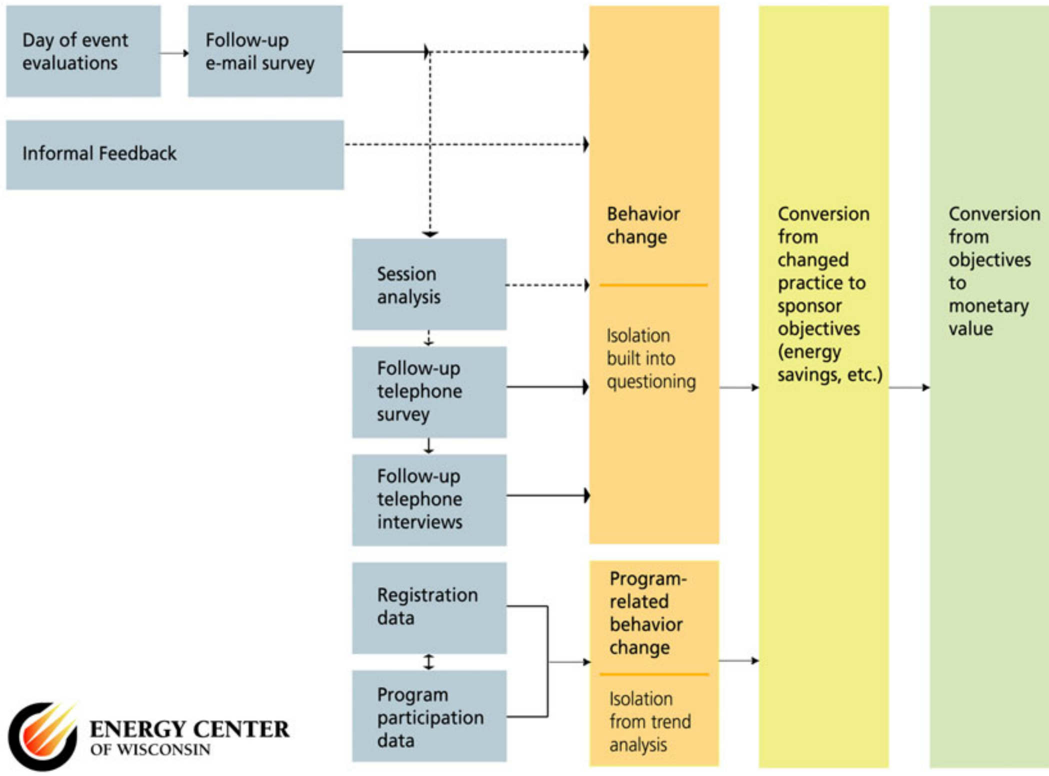
The structured interviews – which were guided by a computer-aided telephone interviewing system – first identified the aspects of new home construction (or remodeling) in which the interviewee was involved (e.g., framing, insulation, HVAC specification). We created nine residential-construction related categories in all (plus a catch-all for respondents who might not fit elsewhere on the list). The system then cycled the interview through a series of questions for each of the categories into which the respondent self-identified. The first question asked "Did you learn or hear about anything new at the conference that prompted you to do something different in . . . ?" If the respondent answered affirmatively, the system prompted the interviewer to ask "What are you doing differently?" and "How is this different from what you did before?" – and record the open-ended responses. The system also prompted to find out the fraction of homes to which the stated change applied. (Elsewhere in the interview we obtained an estimate of annual homes built or remodeled per year by the respondent's firm.)

Altogether, the 24 interview respondents included in our analysis self-reported 29 changes in practices that they ascribed to having attended the conference: 9 of 16 interviewed builders reported at least one change in practice, as did 3 of 4 interviewed remodeling contractors and 3 of 4 subcontractors.

[3] The number of interviews with remodelers exceeds our sampled number of such businesses because some firms we categorized as builders self-identified as remodelers during the interviews.

[4] We did not classify any businesses as "other," but interview responses caused us to classify three respondents as "other" subsequently.

Figure 1. Approach to Data Collection and Analysis



We also asked several questions intended to provide context and allow us to assess the likelihood that the conference had actually caused behavior change. These questions included two in which we asked the respondent to assess the effectiveness of the conference and the sponsoring program overall in improving those aspects of homes that the Wisconsin ENERGY STAR programs are trying to affect. (Energy efficiency is one of these.)

Estimating First Year Energy Savings

We next estimated a range of annual energy savings for each reported change in practices. Though we hypothesized – and received reports of – changes in practices that had benefits to housing durability, health, and safety, we opted to focus solely on reported changes involving energy savings.

Depending on the nature of the response, some of these were more speculative than others. Among practice changes with more readily quantifiable impacts were changing from 16-inch to 24-inch stud spacing and specifying electrically-efficient furnaces more frequently: these measures are discrete and have savings that can be reasonably estimated from existing studies or engineering calculations. On the more speculative end were statements such as "I'm doing a lot better job of air sealing," and "using different kinds of insulation." For these, we tried to estimate a likely range using our best professional judgment – and erring on the side of being conservative. For example, we assumed that "doing a lot better job of air sealing," might result in a 200 to 500 cfm_{50} reduction in blower-door measured leakage, which would translate into 15 to 25 therms of natural gas saved per year, or 2.5 to 3.5 percent of the typical 700-therm heating usage for a new Wisconsin home.

For some reported changes in behavior, it was not clear whether any energy savings would result. ("Using different kinds of insulation" was one example.) We did not include these. There were also a couple of cases in which a reported change in practice could increase energy consumption: we included these as negative savings.

We estimated most of the reported changes to have a small impact in a given home, but a few attendees reported larger changes (Table 2).

Table 2. Sample Measures Attributed to Conference and Energy Savings Assigned

Measure Reported	Number of Respondents	Range of Energy Savings (per year per home)
Small Measures		
increased efforts on air sealing	8	2 to 25 therms
changed framing technique	4	0 to 30 therms
insulation changes	5	-30 to 25 therms
increased ENERGY STAR lighting	3	50 to 375 kWh
Large Measures		
specifying geothermal for some homes	2	625 to 725 therms -9000 to -5000 kWh
increased specification of electronically commutated motors in furnaces	2	700 to 1400 kWh

Across our interviewed builders, remodelers, and subcontractors, these measures suggest a blended average first-year energy impact of 100 to 230 therms and 230 to 700 kilowatt-hours per firm[5]. (As a result, our model suggests a total first-year energy impact of the conference of 13,000 to 31,000 therms and 30,000 to 93,000 kilowatt-hours by Wisconsin-based attendees.)

Extrapolating Results to Measure Life, Persistence, and Population of Attendees

The last step of the process was to estimate the total impact of the conference in terms of energy savings. This involved several additional sub-steps: (1) extrapolate the annual savings to the lifetime of the measures, (2) estimate the persistence of change in practice over time, and (3) extrapolate our interview sample to the conference as a whole. We implemented all of these (along with the uncertainty in first-year savings) within a probabilistic analysis known as Monte Carlo analysis. Monte Carlo analysis is an iterative technique that propagates uncertainty in various input parameters (such as first-year savings and measure life) through an analysis to provide an estimate of uncertainty in outputs (i.e., the total savings attributable to the conference). Given the considerable uncertainty involved in several steps of our analysis, we did not want to ascribe a false level of precision to the results.

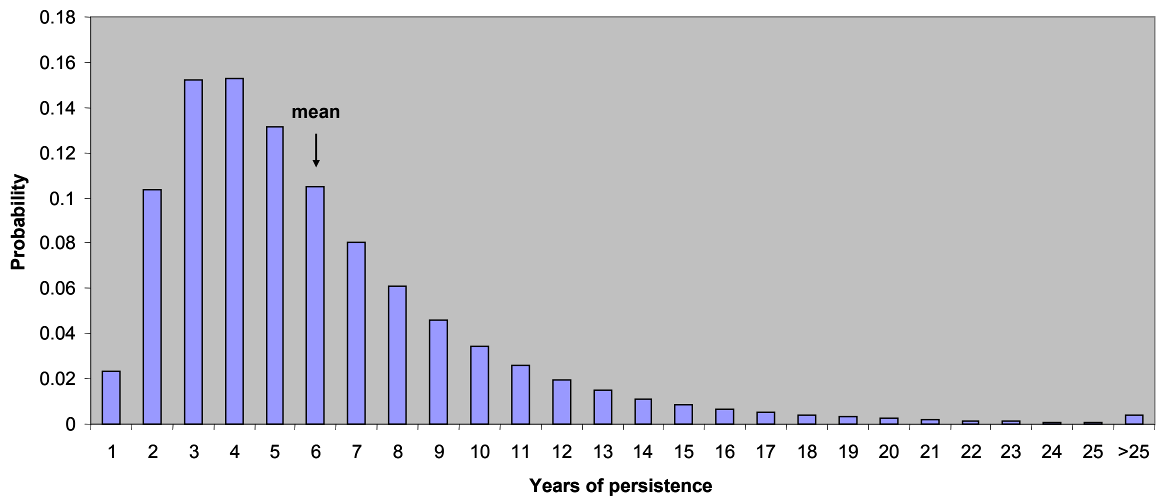
We used three ranges for measure lifetimes: 20 to 40 years for practice changes that affect the building shell and therefore are relatively permanent; 15 to 25 years for changes involving equipment such as furnaces that have an expected lifetime in this range; and, 2 to 7 years for lighting-related changes. The last may be conservative insofar as the conference

[5] The average number of homes affected per firm is between 32 and 44 (12-19 for builders, 48-110, and 57-85 for subcontractors).

prompted attendees to specify hard-wired fluorescent fixtures that will continue to be fluorescent even after the initial bulbs burn out.

Gauging the persistence of practice changes (i.e., how long a builder will continue to implement a given change in practice in the homes he or she builds or remodels – or how long it would be before he or she would have changed practices in the absence of the conference) is perhaps the most speculative aspect of the analysis. We speculated that the net duration in practice changes would in many cases be fairly short, but that it was possible that a young builder could pick up something from the conference that would have a very long net persistence. We therefore defined a log-normal probability function for duration, as shown in Figure 2. In the Monte Carlo analysis, each practice change is randomly assigned a persistence according to this distribution: this results in most practice changes being assigned a duration of less than 10 years, but still allow for the possibility of an occasional long-lasting change.

Figure 2. Estimated Persistence of Practice Changes



For a given change in practice, the total impacts are then calculated as:

Total savings = estimated annual energy savings per home * reported homes built/remodeled per year * reported fraction affected by change in practice * estimated duration of change in practice * measure life

The estimated uncertainty in each portion of the equation above is reflected in the Monte Carlo analysis as random fluctuation (within our assigned uncertainty range) in the values of the inputs for each reported change in practice.

To monetize energy savings and discount future benefits to the present, we used estimated Wisconsin utility avoided costs of 4 to 8 cents per kilowatt-hour and $0.60 to $1.40 per therm (Pigg et al. 2005), and a discount rate of 5 to 15 percent. Because these inputs were included in the Monte Carlo analysis as ranges, the results we present later also reflect uncertainty in avoided costs and discount factor.

Finally, we needed to extrapolate from our interview sample to the full population of builders, remodelers and subcontractors who attended the conference. We did this by randomly assigning an interview respondent – including those who reported no changes – to represent each firm in the un-interviewed population of attendees for each Monte Carlo iteration, then aggregating results across the population. This assignment was done with replacement (and within the categories of builder, remodeler and subcontractor), so that a given interview respondent could be randomly assigned to multiple firms in the un-interviewed population of firms. Over the 5,000 Monte Carlo iterations that we ran, the net result of this approach is to both extrapolate the interview sample to the full population and also incorporate the uncertainty in the interview sampling process itself. It requires assuming only that the sample of interviewees is reasonably representative of the population of firms, an assumption that appears tenable based on the information available to us.

Study Results

Although our estimated energy savings for any one house are small, the aggregate impact of the conference involves extrapolating across an average estimated measure life of more than 20 years, an average estimated persistence of the practice of six years, and across the multiple homes built or remodeled by each firm (which ranged from 1 to 200 across the interview respondents).

When fully run out in this way, the results indicate a societal present value of somewhere between $400,000 and $2,000,000 (representing the 90% confidence range from the Monte Carlo analysis), compared to an overall conference cost of about $380,000. (The conference cost is based on all labor and direct costs involved with the planning and production of the conference.) This represents a benefit/cost ratio range of somewhere between about 1.1 and 5.3, or a return on investment of about 5 to 425 percent. The Monte Carlo model suggests a 97 percent probability that the present value of the future energy savings exceeds the conference costs.

If there are positive, quantifiable benefits from any of the other hypotheses or from out-of-state attendees, actual returns to sponsors are higher. The data available to us suggests that there are such benefits. For example, survey respondents recalled two sponsors' financial contributions to the conference in response to an open-ended (unaided) question even though we implemented the survey ten months after the event. For most sponsors, there is value in the lasting connection attendees made between the conference and the sponsor.

Determining the return on investment for <u>individual</u> sponsors poses some additional challenges because benefits could be allocated to the sponsors in different ways and there is no direct linkage between sponsorship dollars and reported behavior changes. For example, the Wisconsin ENERGY STAR programs are the only conference sponsors currently being evaluated based on energy savings. If one were to assign all energy savings to these sponsors, the program's return for its sponsorship would be 740 to 3,900 percent. On the other hand, if one assigned energy savings proportionately to all sponsors, the return would be 5 to 425 percent. Neither of these extremes feels particularly satisfying: The low set of numbers would leave much energy savings without allocations to anyone because the other sponsors would be unlikely to seek credit for it. The higher set of numbers implies causality where it probably did not exist. How these savings should be allocated is still an unresolved question.

Moving from Internal Study to Replicable, Evaluator-Accepted Approach

This study was a relatively low-cost internal impact assessment of an Energy Center conference, but the approach we followed also could be used for external evaluations of energy efficiency programs that feature or include training components. The wide ranges of our impact estimates – while acceptable for our purposes because they clearly suggest an overall positive return on investment – may prove problematic for evaluators who need to report annual impact estimates. We suggest the following adjustments and enhancements to reduce the uncertainties in a study like this.

Gather More Detail about the Behavior Change Implemented

We trained administrative staff to conduct the telephone interviews using a highly structured computer-aided survey instrument. While we think our interviewer did a very good job, she was not able to ask some follow-up questions that a building expert could devise during an interview. In the future, we would include a follow-up process whereby a technical expert contacts the respondent about ambiguous or vague descriptions of behavior change.

Ask about Implementation Caused by Multiple Interventions

Our survey was clearly about the conference, which holds the potential for socially desired response bias. We were aware of this risk and attempted to account for it when assigning energy savings, especially when responses to other questions seemed inconsistent with strong impacts from the conference. One way to reduce respondents' overstatement of causality is to ask questions in the context of multiple program measures (i.e., ask the extent to which various potential causes contributed to an implementation of a new idea).

Move Beyond Self-Reports

We relied exclusively on self-reports by participants and a dose of conservative skepticism built into our analysis to determine what was implemented as a result of the conference and how effectively it might have been installed or implemented. Having an experienced residential building researcher estimate the energy savings was a critical part of our study. However, in the long run – and given more financial resources – we think it would enhance the credibility of results to move beyond self-reports to ensure that we fully understand what changes have been made and whether they are done in a way that is likely to produce the potential energy impacts. These could include site inspections, blower door tests, and perhaps even billing data analysis or metering studies if reported changes are substantial enough to warrant employing these impact evaluation tools.

Track Persistence

Lack of empirical information forced us to rely on an educated guess about the persistence of the conference-induced behavior change and assign a wide uncertainty around that guess. In the case of a training-induced behavior change, persistence is a function of the number of years before a training attendee (1) will give up the newly learned practice, (2) will cease the activity in which he or she applies the practice (e.g., stops building houses), or (3) would have adopted the practice for another reason. More data on any of these issues would allow a more educated estimate of persistence. Ideally, we would conduct follow-up surveys of attendees and study developments in professional practices followed by the attendees' peers. Even in the absence of such data, knowing the approximate ages of attendees and typical years that attendees stay in their current professional field would be helpful.

Resolve Allocation Issues When Multiple Factors Might Take Credit

As noted above, we face a challenge in allocating the impacts we identified to the various conference sponsors. While this situation is unique to multi-sponsor events, a similar allocation challenge can occur when multiple program elements and other market forces all work together to cause a change in practice. The energy efficiency industry may need more discussion about the most appropriate ways to allocate causality and identified impacts back to the various program elements when multiple interventions compete for the same successes. A winner-take-all approach is probably too simplistic, and the programmatic element that has traditionally gotten the credit might not always be the right one or the only one.

Implication for Energy Efficiency Evaluations

We think the results of this modest effort reinforce the notion that training impacts have the potential to be significant and should not be overlooked by evaluators – especially when program funding and design gravitates to those program approaches that show measurable energy impacts. The art of measuring training impacts on energy efficiency is not nearly as well-developed as many other aspects of energy evaluation, so all attempts to measure training impacts add to our collective understanding of how training events contribute to energy savings. In this context, it is important to note that education and training approaches tend to be much more varied than other efforts to stimulate energy efficiency, and not all approaches are designed to produce demonstrable changes in behavior. Nonetheless, it would serve the industry well if more attempts were made to estimate energy savings from education and training events, and the methods and results were shared publicly. Peer review of these studies is critical to build industry consensus on the minimum requirements and best practices to achieve credible results that will be accepted first by the energy evaluation community and subsequently by program funders, program designers, regulators, and policy-makers.

References

Eley, C. 2004. "Raising the Bar: How California's New 2005 Standard 478 GWh and 181 MW a Year." In *Proceedings of the ACEEE 2004 Summer Study on Energy Efficiency in Buildings*, 4:84-94. Washington, D.C.: American Council for an Energy-Efficiency Economy.

Green, J., and L. Skumatz. 2000. "Evaluating the Impacts of Education / Outreach Programs: Lessons on Impacts, Methods, and Optimal Education." In *Proceedings of the ACEEE 2000 Summer Study on Energy Efficiency in Buildings*, 8:123-136. Washington, D.C.: American Council for an Energy-Efficiency Economy.

Kirkpatrick, D. 1994. *Evaluating Training Programs: The Four Levels*. San Francisco, CA: Berrett-Koehler Publishers, Inc.

Phillips, J. 2003. *Return on Investment in Training and Performance Improvement Programs*, Second Edition (Improving Human Performance). Butterworth-Heinemann Publishers, Inc.

Pigg, S. et al. *Energy Efficiency and Customer-Sited Renewable Energy: Achievable Potential in Wisconsin 2006-2015*. Madison, Wis.: Energy Center of Wisconsin.

RLW Analytics. 2005. *Impact and Process Evaluation: Building Operator Training and Certification Program*. Lexington, Mass.: Northeast Energy Efficiency Partnerships, Inc. (Report available at http://www.efficiencymaine.com/pdf/BOCfinalreportdelivered.pdf.)

Lawrence Berkeley National Laboratory and Xenergy, Inc. 2004. *Evaluation of the Compressed Air Challenge® Training Program*. Washington, D.C.: U.S. Department of Energy, Energy Efficiency and Renewable Energy. (Report available at http://www.nrel.gov/docs/fy04osti/35301.pdf.)

York, D. et al. 2002. "Administration and Implementation of Public Benefits Programs: Experiences from Four States." In *Proceedings of the ACEEE 2002 Summer Study on Energy Efficiency in Buildings*, 5:327-346. Washington, D.C.: American Council for an Energy-Efficiency Economy.