How WISE Evaluation Can Improve Contractor Training for Home Electrification

Jen Loomis, PhD, Opinion Dynamics
Ellen Steiner, PhD, Opinion Dynamics
Erin Kempster, Opinion Dynamics

ABSTRACT

Whole Independent Systems Evaluation™ (WISE™) is an evaluation approach that maximizes timely insights to allow for course corrections in program implementation. Our firm has applied WISE principles to workforce, education, and training activities designed to improve contractors' knowledge, skills, and confidence working with heat pump equipment. The market for heat pumps in California is in an early stage, and ensuring customer satisfaction is necessary to advance the market.

Using models for evaluating adult learning and the Technology, Program, Market, and Policy Watch pillars that form the basis of WISE, we were able to provide timely feedback to program administrators, who used it to make course corrections that will improve contractor and customer outcomes and increase the adoption of heat pump technology.

In our evaluation, we identified issues in the way heat pump trainings were offered to contractors. We also recommended improvements to the way the program administrators measured adult learning outcomes from heat pump trainings. Our customer surveys also uncovered installation issues that can be corrected via proper training. The program administrators found the feedback valuable and took corrective steps. This paper explains how the WISE approach facilitates the timely identification of issues and the appropriate changes that allow implementers to maximize the learning outcomes of trainees and allow for improved evaluability.

Introduction

Heat pumps are the technology of choice for a growing number of residential programs aimed at improving energy efficiency or reducing greenhouse gas emissions. Despite a well-developed market for heat pumps in Europe and Asia, heat pumps are a relatively new technology for some US markets, including California. The State of California and organizations within it have dedicated hundreds of millions of dollars to implement heat pump market transformation programs.

When considering how to assess market transformation for innovative technologies such as heat pumps, Everett Roger's Diffusion of Innovation model (2003) provides a solid foundation. This adoption model framework characterizes where a technology's market is in the adoption curve and what market metrics are most important to track given the market adoption stage. Rogers' Diffusion of Innovation model describes stages of market adoption through a progression in the types of consumers that the technology will likely attract throughout its adoption lifecycle. This classification includes five adopter stages, from innovators to laggards.

Geoffrey Moore, in his book Crossing the Chasm (2014), states that the most difficult step for emerging technology is making the transition between early adopters and the early majority (Figure 1). Innovators and early adopters are dominated by a small group of visionary

customers who are quick to appreciate the nature and benefits of the new technology, whereas the early majority represents the mainstream market who are largely pragmatists in nature. Heat pumps in California may be described as being in the early adopter stage.

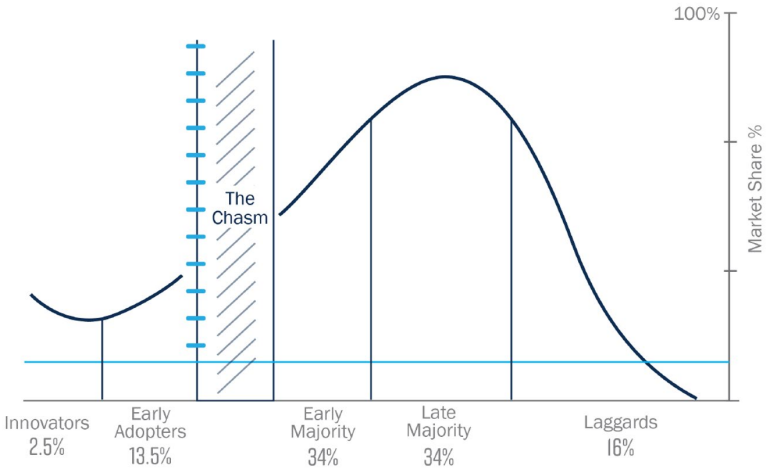

Figure 1. Technology Adoption Life Cycle Model

In our significant experience evaluating market transformation programs across the country, we have observed that workforce and program implementation strategies are critical to successful technology adoption programs that support a technology effectively "crossing this chasm." The workforce needs to be aware of and knowledgeable about technology such as heat pump water heaters; they also need the skills to install them and realize the value of selling and promoting them. At this early stage in the marketplace, the role of the installer is paramount to market adoption.

In the residential market, installers are the voice for the technology, and they frame the perceptions of the technology for customers. Given that space and water heater purchases occur every 10 to 15 years, installers often act as the "trusted advisor" in making this infrequent purchase decision. To make the jump from early adopters to early majority, market transformation initiatives must consider not just the consumer diffusion curve but also the workforce diffusion curve.

For HVAC and water heating workforce education and training (WE&T), the number of attendees is an easy metric to collect and report, but the number of people who attend training is merely an output and does not predict learning outcomes. As heat pump technology has improved and reliance on that technology as a decarbonization pathway has increased, it has become ever more important that installation contractors are provided with knowledge and resources that are accessible, intelligible, and actionable so that they can go forward from such trainings with confidence and skill that resonate in the real world.

Establishing WE&T frameworks and content that resonate within the day-to-day lived experiences of contractors requires more than bringing on experts with good intentions and relevant knowledge. To ensure that training participants fully comprehend the material and are able to translate that understanding into action requires thoughtful construction of a training paradigm built upon established frameworks. To fully optimize learning outcomes, a training plan must feature two key components: First, evaluation happening concurrently with program implementation, and second, feedback mechanisms from the field to the training providers.

Whole Independent Systems Evaluation™

Whole Independent Systems Evaluation, or WISE™, is an adaptation of the long-standing practice of Developmental Evaluation and has been tailored to suit the needs of modern decarbonization and electrification efforts.[1] As such, WISE utilizes a combination of traditional process, impact, and market transformation evaluations. However, it takes a proactive approach to evaluation by generating insights from the very beginning instead of waiting until the end. This allows for immediate feedback and optimization of outcomes throughout the implementation process.

To ground oneself in the concepts of Developmental Evaluation, it is helpful to review the writings of Michael Quinn Patton, who is widely considered to be the progenitor of this evaluation model. As explained by Patton, "…because evaluation typically carries connotations of narrowly measuring predetermined outcomes achieved through a linear cause-effect intervention, we want to operationalize evaluative thinking in support of social innovation through an approach we call developmental evaluation. Developmental evaluation is designed to be congruent with and nurture developmental, emergent, innovative, and transformative processes" (Patton 2006:28). Or, as summarized by an attendee of an American Evaluation Association event in 2015, developmental evaluation "allows evaluators AND program implementers to adapt to changing contexts and respond to real events that can and should impact the direction of the work." (Parkhurst et al., 2016)

As traditional energy efficiency and resource acquisition programs began to transition into decarbonization and electrification programs, we recognized a similar need to transition evaluative thinking from a summative approach to a model that better reflects and reinforces the objective of making decisions in uncertain and complex contexts. In undertaking this shift in thinking, Developmental Evaluation offers a useful heuristic from which to draw and reflect on the particular challenges and dynamics of the clean energy industry.

Developmental Evaluation is organized along eight principles that emphasize the timeliness and utilization of feedback in a dynamic context (Patton, McKegg, Wehipeihana 2016). In customizing the framework to enable industry-specific program optimization without delay, we retained and refined many of the core tenets to best meet the needs of the clean energy economy. The result was the set of six UTOPIA principles of the WISE model as follows:

- **U**tilization-Focused: Actionable insights enable program iteration and optimized results;
- **T**imely: Concurrent monitoring facilitates course correction and progress without pause;
- Systems **O**utlook: Embeds analysis within the larger ecosystem context;
- Complexity **P**erspective: Methodologies reflect interconnected dynamics of baseline and emerging conditions;
- **I**ndependent: Independent, implementer-agnostic third-party evaluation oriented towards outcomes and with no vested interests; and
- **A**daptive: Leading insight delivery through changing conditions.

Another hallmark of developmental evaluation is that it "does not rely on or advocate any particular evaluation method, design, tool, or inquiry framework" (Patton, McKegg, Wehipeihana 2016). In recognition of the myriad factors at play that can speed or slow market adoption of new clean energy technologies, we determined the importance of focusing

[1] The Four Pillars of Whole Independent Systems Evaluation (WISE) with Market Watch, Program Watch, Technology Watch, and Policy Watch is a trademark of Opinion Dynamics.

observational and evaluation activity along four key pillars: Market Watch, Program Watch, Policy Watch, and Technology Watch.

- **Market Watch** identifies, tracks, and measures changes to the diffusion of technology adoption, changes to barriers, and shifts in the supply chain as program activities address barriers over time;
- **Program Watch** provides fast feedback to implementers and assesses scalability potential. The focus of this pillar is on evaluating the effectiveness of a program;
- **Technology Watch** monitors technology development, performance, and pricing; and
- **Policy Watch** documents and tracks changes to external incentives and policy decisions at the utility, state, and federal levels.

Any program's WE&T activities touch upon the market, program, and technology pillars. As technology changes, it becomes all the more important for a program to adapt its workforce training content, the effectiveness of which can be observed within the market watch pillar.

WE&T-Specific Evaluation Framework

Uncovering the barriers and conduits to optimal outcomes from WE&T initiatives requires the use of purpose-built frameworks. This is especially important to truly understand the causal links between installer training and installer practice. For example, quality HVAC repair and installation is a highly technical activity in which improper execution of the necessary steps can lead to incorrect diagnoses of problems as well as solutions that can decrease efficiency. Thus, it is essential that trainings emphasize proper installation practices for installers to maximize energy savings. Past research has demonstrated that market actors who participate in WE&T activities often make changes to their practices that result in energy savings and that a large majority share what they learned with others (Opinion Dynamics et al., 2018). Thus, the training course impacts extend beyond the attendee and makes the potential for savings quite large. Yet, it is important to ascertain causal links between installer training and installer practice.

WISE research in this area is guided by Donald Kirkpatrick's Model for training evaluation. Kirkpatrick's conceptual framework addresses the overall assessment of training programs. The Kirkpatrick Model is the gold standard framework in adult training circles for assessing training programs (Kirkpatrick and Kirkpatrick 2006) and is noted in The California Evaluation Framework discussion of Information and Education programs. As illustrated in Figure 2, Kirkpatrick's Framework consists of four levels.

The first level is Reaction. Reaction measures how participants feel about the learning experience. The value of Level 1 is that a good training experience improves knowledge transfer. Level 2 is Learning. Learning assesses the degree participants change attitudes, increase knowledge, or enhance skills as a result of the learning experience. The value of Level 2 is to demonstrate that learning occurs as a result of the training. The third level is Behavior. Behavior measures the degree to which participants apply what they have learned outside of the learning environment. This level seeks to demonstrate whether trainees take the information they learn and apply it. Finally, Level 4 is Results. Results refer to the degree to which targeted outcomes are achieved system-wide. For results, we seek to measure the program's overall impacts and tangible results, such as energy savings, job creation, job placement, improved quality, and increased productivity. The value of measuring Level 4 is to inform the return on training investment that a program, entity, or organization realizes from the training endeavor.

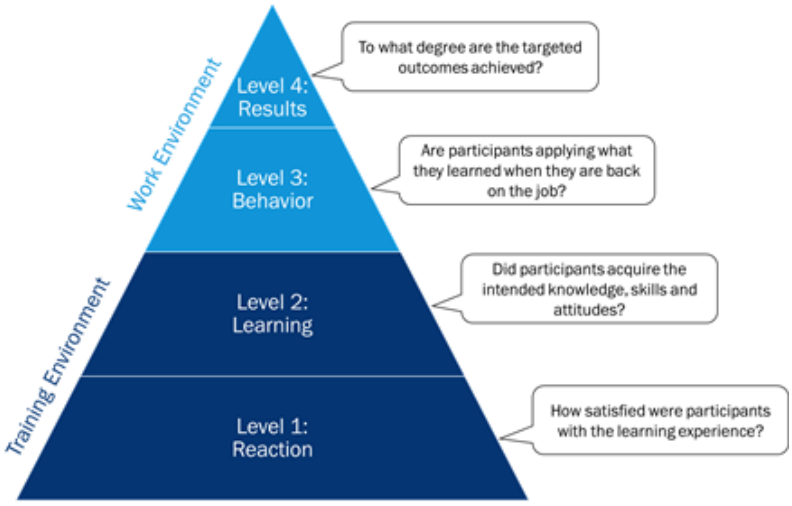

Figure 2. Kirkpatrick's Learning Framework

In undertaking any new initiative that incorporates WE&T, it behooves the program administrator to contract for evaluation services at the same time as implementation services to enable training plan optimization ahead of delivery, as well as to facilitate pulse checks and enable rapid refinement as implementers operationalize that plan. Enabling this rapid refinement is the incorporation of model-driven queries into any surveys of training participants.

For instance, questions aligned with Kirkpatrick's Level 1 may cover the extent to which training materials are purely academic versus translatable to the contractor's day-to-day work; whether they enable the trainee to apply what they have learned; and whether the level of material is calibrated to the learners' levels of expertise.

Inquiry in-line with Level 2 measures the degree to which participants change attitudes, increase knowledge, and enhance skills as a result of the learning experience. Tools for this inquiry may involve pre- and post- tests administered via web survey to measure knowledge and skills gain. Another option is to use simulations to see the learners' ability to apply the newly gained knowledge and skills in real-world settings.

These same interventions can be used to assess Level 3. Ride-alongs with previous training participants enable further insight, although it may be important to explain to participants that such ride-alongs are focused on refining the curriculum for the class they participated in to reflect more real-world experience as opposed to assessing their installation knowledge, skills, and abilities. Depending upon the objectives for the WE&T effort, ascertaining Level 4 insights may take several forms of evaluation.

Results: Insights and Iterative Improvement

Since 2021, Opinion Dynamics has had the opportunity to apply WISE to a large-scale market transformation program focused on space conditioning and water-heating heat pump technologies. Our evaluation activities touched on the Technology, Market, and Program Watch pillars and allowed for timely feedback to program implementers.

Example 1, Program Watch:

We identified issues in the way heat pump trainings were marketed and communicated to contractors. Early on in the program, it was challenging to get a full schedule of training events. After the program administrator provided a schedule, it quickly became out of date when we found that upcoming trainings had been canceled due to low registration. We also observed that different lists of available trainings were housed in a multitude of places. While this can be advantageous for outreach and awareness, there was not one site that clearly listed all of the program-sponsored trainings. As such, the uncertainty about the training offerings made it difficult for potential attendees to plan their schedules and take time off work to attend them.

We also noticed that all of the trainings were presented side-by-side with no guidance as to which one contractors should take first. For example, four different HVAC heat pump courses were offered at the same time. Effective training builds upon what was learned in prior courses, allowing the attendee to advance their knowledge deliberately. The TECH trainings, on the other hand, took a scattershot approach and offered the contractor no clear starting and ending point.

Through surveys with contractors, Opinion Dynamics showed that nearly a quarter of contractors had not heard of the free trainings offered through the program (38 of 184; 21%). Among contractors who had learned of the trainings, a minority attended them (54 of 168; 37%). The WISE approach enabled Opinion Dynamics to provide these findings to the program staff early on. The program staff took corrective action and revamped a contractor-facing website. The website now has an identifiable training hub that houses all available contractor trainings in one place. It also has a calendar and clearly organized schedule that makes it easier to see when the trainings will occur.

Example 2, Program Watch:

At the outset of a new training offering on heat pump water heaters, the program administrator shared with us their plans for a post-training survey. We identified that the survey was in alignment with Level 2 of Kirkpatrick's model to demonstrate an increase in knowledge or enhanced skills as a result of the learning experience. The program staff, however, had not planned to conduct pre-tests of knowledge and skills. We recognized the importance of having a baseline measurement from which to judge increases in knowledge caused by the training. In WE&T, it is impossible to go back and create a baseline measurement because the learner has already been exposed to the new information. Therefore, there is only one chance to conduct a pre-test and capture baseline knowledge. Given our position as WISE evaluators and engaging with the program administrator before they issued the survey, we were able to advise them on how best to construct their data collection forms to accurately capture the impacts of the training. Without our recommendations, the implementer risked either overstating or understating the training's impacts.

We recommended they make two important changes. The first was to conduct a pre-test to capture the level of knowledge and skills of the attendees prior to the training. That way, the implementer can calculate the increase in knowledge provided by the training and avoid the risk of overstating the training's impacts. The second change was to adjust some questions in the post-test. The post-test questions assumed knowledge gain beyond what could reasonably be achieved in a one-hour webinar. We recommended questions that would more appropriately capture what someone could expect to learn in that amount of time in an online group setting.

By setting reasonable expectations for the learning outcomes, the implementer avoided the risk of understating the training's impacts.

Example 3, Technology and Program Watch:

In order to learn the effectiveness of the program's WE&T activities in line with Level 4 of the Kirkpatrick model, we conducted a survey one year into program operation with customers who received an incentivized heat pump at least six months prior. The nearly 1,000 customer responses yielded valuable learnings that, in the spirit of the UTOPIA principles, we were able to provide to the program implementers. Additionally, insights into various negative experiences with the installed equipment led our team to make recommendations specific to the WE&T courses' curriculum. This effort supported our Technology Watch by uncovering customer impacts derived from the heat pump technology, and informed Program Watch by informing the Program's contractor training activities.

In our study, we learned that about one in five heat pump water heater (HPWH) customers (56 of 300; 19%) needed to repair, replace, or troubleshoot issues with their newly-installed HPWHs. And, more than 100 of the 300 noticed issues with noise or vibration or other installation issues that point to contractor training opportunities. Of customers who reported vibration and noise issues, a large majority were at least slightly bothered by them (Figure 3). Customers noticed noise more when the HPWH was installed in conditioned spaces inside the home. HPWH noises seemed to both people most when it disrupted their sleep. One customer wrote in our survey, "The noise level is so loud my kids can't sleep sometimes. I had to adjust the time when it heats up." In nine cases, the contractor returned to install a foam kit or vibration isolation kit designed to reduce noise and vibration.

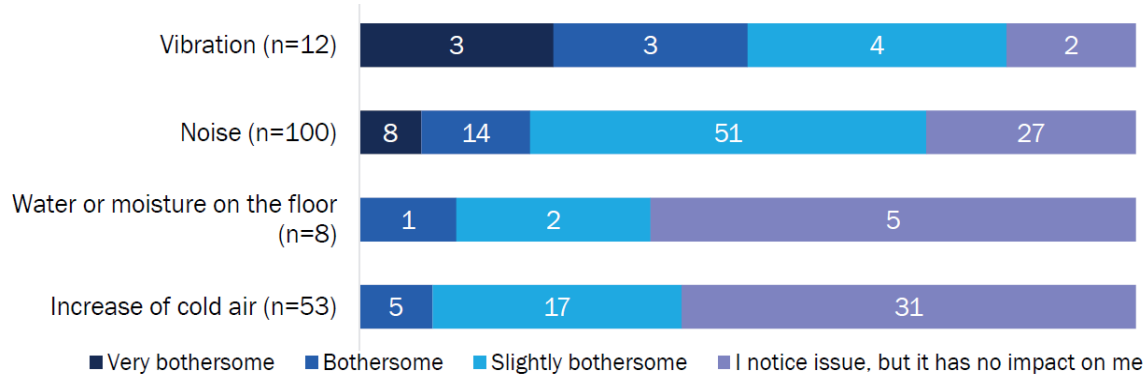

Figure 3. Extent to Which HPWH Issues Bothered Customers

These findings allowed us to share with the program implementer that the training organizations they have partnered with should include instruction on how to avoid noise and vibration issues as well as leaks. Specifically, they should focus on installation practices that reduce noise and vibration issues such as not installing the HPWH near a bedroom wall or installing a vibration isolation kit.

Our survey also showed that 60 of the 494 ducted HVAC customers noticed that the air coming out of the vents was not as hot when heating the home with their prior system, which bothered 37 of them (62%). The performance of HVAC heat pumps is sufficiently different from furnaces such that customers should be advised about what to expect in terms of run times and

the air temperature coming out of the vents. Contractors should educate their customers about these differences during the sales process to adjust their expectations in advance. By providing customers with realistic expectations, the number of callbacks will likely decrease and lead to a more satisfied customer. Again, this finding highlights an opportunity to optimize the WE&T course offerings available to the program's installation contractors.

Many issues we encountered in the survey are largely avoidable with proper training and execution of industry best installation practices. Once contractors are trained on the installation issues, it will lead to greater customer satisfaction. Customer satisfaction will lead to more positive word-of-mouth about heat pump technologies and recommendations for the equipment. Positive customer testimonials are critical right now, at a crucial juncture in the technology's adoption curve. Customer support is needed to move heat pumps along the adoption curve across the chasm from early adopters to the early majority.

Conclusion

Greater heat pump adoption is needed to meet state and utility goals. Heat pumps are at a challenging point on the diffusion curve and will need all of the positive customer referrals they can possibly get. Our research with heat pump customers has revealed there are some aspects of heat pump performance that are bothersome to customers. The issues can be remedied through improved contractor training.

Applying the WISE evaluation approach to a heat pump market transformation program generates tremendous value. The evaluator was able to assess the promotion of the WE&T offerings and recommend improvements to the way the schedule was organized and communicated. The evaluator was also able to conduct surveys with participants at key points during or after their participation to uncover issues and generate timely insights. By using the Market Watch, Technology Watch, and Policy Watch, the evaluator can offer the most pertinent and actionable recommendations to the program. This gives the program implementers the chance to course correct to make iterative improvements while the program is still running.

In this case, applying the WISE framework enabled Opinion Dynamics to provide advice that will ensure the WE&T offerings cover the most important information to ensure optimal heat pump performance and high customer satisfaction. Instead of the program administrator having to institute a new WE&T effort to mitigate these issues in the market post-program, the WISE-generated feedback to the in-flight program enables iterative improvements along the way. The result should lead to an increase in properly installed heat pumps, fewer equipment issues, greater customer satisfaction, and greater adoption of heat pumps.

When the program administrator and training organizations follow the advice generated from WISE evaluation, it increases the likelihood that heat pumps will move along the adoption curve and soon become the equipment of choice for both consumers and contractors. Our next step will be to follow up with the program implementer to see how successfully the training organizations have been able to incorporate updates into their curriculum to focus on the biggest problems reported by customers.

References

Kirkpatrick, Donald and James D. Kirkpatrick. 2006. *Evaluating Training Programs: The Four Levels.* San Francisco: Berrett-Koehler Publishers, Inc.

Moore, Geoffrey. 2014. *Crossing the Chasm: Marketing and Selling Disruptive Products to Mainstream Customers*. New York: Harper Business Essentials.

Opinion Dynamics, Wirtshafter Associates, Inc., Jai J. Mitchell Analytics, Summit Blue Consulting, "Indirect Impact Evaluation of the Statewide Energy Efficiency Education and Training Program." Accessed May 5, 2018 at http://www.calmac.org/publications/06%2D08%5FStatewide%5FEducation%5Fand%5FTraining%5FImpact%5FEval%5FVol%5FI%5FFINAL%2Epdf

Parkhurst, Marcie, Hallie Preskill, Jewlya Lynn, and Marah Moore. 2016. *The Case for Developmental Evaluation*. Accessed March 6, 2024 at https://www.fsg.org/blog/case-developmental-evaluation/

Patton, Michael Quinn. 2006. "Evaluation for the Way We Work." *The Nonprofit Quarterly*. Vol. 13 (1): 28-33. Accessed March 6, 2024 at http://www.scribd.com/doc/8233067/Michael-Quinn-Patton-Developmental-Evaluation-2006

Patton, Michael Quinn, Kate McKegg, and Nan Wehipeihana. 2016. *Developmental Evaluation Exemplars: Principles in Practice.* New York: The Guilford Press.

Rogers, Everett M. 2003. *Diffusion of Innovations*. 5th Edition. New York: Free Press.