July 6, 2023

REPORT #E23-468

Assessment of NEEA's Approach to the Evaluation of Market Transformation

Prepared For NEEA:
Amy Webb, Sr. Manager of Market Research and Evaluation

Prepared by:
Michael A. Harnar, Ph.D.,
Assistant Professor, Western Michigan University
Owner, Pointed Arrows Consulting

Pointed Arrows Consulting
9934 West Fg Avenue
Kalamazoo, MI 49009

Northwest Energy Efficiency Alliance
PHONE
503-688-5400
EMAIL
info@neea.org

Table of Contents

Executive Summary

In May 2021, the Northwest Energy Efficiency Alliance (NEEA) contracted with Michael Harnar of Pointed Arrows Consulting, to conduct an independent review of NEEA's approach to evaluating its Market Transformation (MT) programs. NEEA's Market Research and Evaluation (MRE) team is the internal functional group that develops the scope of these evaluations, and then manages the contract with the third-party evaluation contractor from project kick off to delivery of the final report. This study reviews MRE's approaches to MT evaluations by interviewing NEEA staff, reviewing program documentation, and interpreting MRE practice through the lens of two evaluation tools: a set of key practices identified in the MT literature and published by NMR (2013) and the Joint Committee on Standards for Educational Evaluation *Program Evaluation Standards* (PrgES). The purpose of the study is to describe MRE evaluation practice as a means to identify strengths and areas of possible improvement for MRE.

Approach

This report is a third-party assessment, or metaevaluation, of the MRE market transformation evaluations. It was conducted by an external consultant, Michael Harnar, and coordinated by Amy Webb, NEEA's Sr. Manager of Market Research and Evaluation. To meet the study objectives, document analysis was completed on 3 different market transformation programs, interviews were conducted with 11 NEEA employees, external sources were found and reviewed for market transformation evaluation knowledge, and finally MRE market transformation evaluation was compared to best practices and an evaluation quality checklist.

Summary of Findings

After reviewing MRE's MT evaluation against the best practices as published by NMR (2013), this report finds extremely laudatory practices. There is regular reflection incorporated into the program so that data and questions are applied to keep focused on outcomes. The theory of change embedded in the market progress indicator tables and the connection of the indicators to the logic model as operationalized outcomes makes transparent the linkages between actions and outcomes and defines expectations of the market. The program theory is established and revisited at regular intervals so that adjustments to MT efforts can be implemented. Though the NMR effective practices model used to review MRE's MT evaluation practices is useful, it skews towards a stronger attribution linkage than what NEEA does.

MRE's MT evaluation scored *very good* when rated on the Joint Committee on Standards for Educational Evaluation (JCSEE) *Program Evaluation Standards* (PrgES). MRE MT evaluation scored highest on propriety, followed by utility, accuracy, feasibility, and finally evaluation accountability. The JCSEE PrgES served as a useful tool to describe MRE's MT evaluation efforts. MRE scored lowest on evaluation accountability, mostly because internal metaevaluation is not made explicit.

Conclusions

NEEA built a smart life cycle timeline to guide decision-making for its Market Transformation programs, called the Initiative Lifecycle, or ILC. The ILC organizes the learning process and reflects the tenets of adaptive management principles. The use of such a learning model, where MRE Scientists are partners in ensuring decision-makers can learn more about their market and make changes when needed, reflects the core of formative evaluation. It becomes summative at key decision moments, such as the decision to move to long-term monitoring and tracking (LTMT). Embodying this iterative learning model is critical in the complex and potentially volatile energy efficiency market.

MRE Scientists are central to the decision-making process that undergirds the MT program development and implementation. They embody the knowledge that goes into so many choices, and *it should be their responsibility to document that understanding*. In late 2020, MRE began relying on a new Research and Evaluation (R&E) Plan template to document the history of the program and the rationale for future areas of research. The R&E Plan template looks like a good start, but ascension to a universally used tool amongst MRE Scientists was contemporaneous with this metaevaluation. This should change and *perhaps the MRE Scientist should take responsibility for using the R&E Plan as a storytelling device that documents the transformation of a market.*

Contribution analysis is a potentially useful way to think about how the preponderance of evidence can be used to tell the MT programming story. As an illustration, if they are diligent to think through all the ways they may be wrong about their claims, gather divergent voices in critiquing their assumptions, then perhaps they can develop a modest claim, through a preponderance of evidence they identified in advance, and be comfortable in that claim.

MRE could improve the credibility of their work by *making explicit its adherence to established evaluation standards and ethical principles*, *working to incorporate more stakeholder voices in the evaluation process*, *exploring plausible alternative explanations to their claims*, and *doing more to document the internal formative evaluation work* that already occurs.

MRE is doing a fine job of being at the table and incorporating evaluative thinking at important moments. Proactive application of communication, negotiating, and methodological skills will be key to providing useful service in the "rough ground" of frontline market transformation evaluation practice.

Introduction

Background on NEEA and Purpose of the Assessment

In the late 1990s, utilities around the Northwest established The Northwest Energy Efficiency Alliance (NEEA) to share the costs and benefits of coordinating efforts to transform markets for energy efficiency. Today, NEEA is a collaboration of 140 utilities and efficiency organizations that work together to advance energy efficiency in the Northwest on behalf of more than 13 million consumers. Since 1997, the region has cost-effectively delivered over 1,700 aMW of energy efficiency through market transformation (NEEA 2020).

NEEA, in collaboration with third party consultants, conducts research and tests hypotheses to identify the barriers that are preventing the adoption of emerging energy efficiency products and practices. NEEA's program teams then work with an array of market partners, including the region's utilities and energy efficiency organizations, but also manufacturers, distributors, end-users, and other market actors, to develop and implement market intervention strategies. Such strategies are intended to remove barriers and exploit market opportunities that accelerate the adoption of cost-effective energy efficiency. The goal of these market intervention strategies is to create lasting market change, or to transform the market. Today, NEEA is considered a national leader in "Market Transformation" (MT) for energy efficiency products and practices (NMR, 2013).

NEEA's MT programs rely on detailed logic models to describe the theory of change associated with each program. To evaluate the effectiveness of its MT programs in achieving logic model outcomes, NEEA uses a mixed method evaluation approach that includes many elements of impact evaluation, as well as process evaluation, in an approach that most closely aligns with "theory-based" evaluations[1],[2]. The Market Research and Evaluation (MRE) department at NEEA is responsible for providing research support and evaluation of the MT programming. Among other evaluation reports that MRE delivers at specific checkpoints to the MT programs they support, a key evaluation deliverable is called a "Market Progress Evaluation Report" or "MPER". MPERs are annual evaluations that track a stable set of market progress indicators (MPIs) which map back to the logic model outcomes. The objective of these MPERs is to provide a "preponderance of evidence" that the MT program is influencing the market in the way posited by its program theory and documented in the logic model. In addition, MRE scopes and manages market research studies, as needed, to refine the program logic. Typical objectives of these studies include better characterizing barriers, honing the value proposition, better understanding the supply chain, and identifying new opportunities.

[1] https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/theory-based-approaches-evaluation-concepts-practices.html
[2] https://www.betterevaluation.org/en/resources/guide/theory-based_approaches_to_evaluation

For some of NEEA's utility stakeholders, who are more accustomed to impact evaluations that quantify energy savings associated with what is referred to in the industry as "resource acquisition", NEEA's theory-based evaluations may be perceived to lack rigor. Whereas utility impact evaluations track quantifiable metrics, in terms of units of a technology rebated[3] multiplied by a unit energy savings (UES) associated with the energy efficiency measure, NEEA's evaluations focus on tracking progress toward market outcomes. These market outcomes are sometimes binary and/or easily observable, such as the adoption of a federal standard. Other outcomes, such as consumer "satisfaction", achieving cost parity with other less efficient products, or a shift in installer rate of recommendation, are more difficult to confirm. Likewise, they can be difficult to attribute to NEEA's market interventions. For this reason, NEEA's evaluations of market transformation programs rely on a preponderance of evidence from multiple sources, often using multiple research methodologies to deliver market intelligence that, in aggregate, "make the case" for market transformation having occurred.

In May 2021, NEEA contracted with Michael Harnar of Pointed Arrows Consulting to complete a third-party assessment of NEEA's general approach to the evaluation of market transformation programs. This assessment has two objectives: 1) to document NEEA's approach to the evaluation of market transformation programs for the benefit of Market Research & Evaluation (MRE) staff and other MRE stakeholder groups (such as NEEA's Cost-Effectiveness Advisory Committee) and 2) to provide specific recommendations, including more appropriate tools, methods, and terminology, to address gaps in the approach and to help MRE better articulate the purpose and mechanics of their approach to the evaluation of market transformation programs.

Throughout this report, the term metaevaluation is often used in place of assessment. While the purposes of this effort are intended to be used for learning and improvement (and some will term that an assessment) the core work involved is meta-evaluative in nature. That is, criteria for what is good about the MRE process and efforts are refined, standards of goodness are set, information is gathered to measure or assess MRE's "goodness", and then some form of judgment is expected. This is the core logic of evaluation and, in turn, the core logic of metaevaluation. Therefore, from here on out, the term metaevaluation is used to describe this project.

While there are likely multiple groups of people interested in this report, it is generally intended to be most useful to MRE leadership as a tool to think about and improve what they do.

[3] Many utility energy efficiency programs offer rebates on approved energy efficiency products, such as energy efficient water heaters, windows, or light bulbs. These rebates are effectively "buying" energy efficiency and offsetting the cost of future demand for energy that would need to be provided by the construction of costly power plants.

Context and Objectives for the Assessment

NEEA solicited Dr. Michael Harnar (the "metaevaluator") to conduct a third-party assessment of its general approach to the evaluation of market transformation programs. Michael Harnar, Ph.D., is the owner of Pointed Arrows Consulting and the Interim Director of the Interdisciplinary PhD in Evaluation program at Western Michigan University. A summary of qualifications for Dr. Harnar is provided in Appendix B. The objectives of the metaevaluation are embedded above, but to be explicit, they are:

1. To document NEEA's approach to the evaluation of market transformation programs for the benefit of MRE staff and other MRE stakeholder groups (such as NEEA's Cost-Effectiveness Advisory Committee). The assessment should identify and document strengths of the approach and areas of alignment with industry best practices, as well as highlight areas for improvement.
2. To provide specific recommendations, including more appropriate tools, methods, and terminology, to address gaps in the approach and to help MRE better articulate the purpose and mechanics of their approach to the evaluation of market transformation programs.

The assessment contained in this report applies a formative metaevaluative approach to serve two purposes: descriptive and evaluative. It is descriptive because it is in the service of evaluation. One needs to develop a deep understanding of something to give it a proper, contextually grounded metaevaluation (Gullickson, 2020; Stake, 1977). It is through describing the MRE's MT program evaluations, using applicable frameworks, that some evaluative insights can be drawn about strengths and potential avenues for improvement. Three questions guide the work in addressing the major objectives.

Guiding Metaevaluation Questions

- **Question 1:** *How do MRE MT evaluation practices compare to industry best practices as framed by NMR (2013)?*

- **Question 2:** *How do MRE MT evaluation practices score on the Joint Committee on Standards for Educational Evaluation Program Evaluation Standards (JCSEE PrgES) checklist?*

- **Question 3:** *What recommendations emerge from the comparison to best practices and the JCSEE PrgES evaluation quality checklist?*

Criteria and Standards for the Assessment

Two major criteria are used in this report. The first is MRE's alignment with best practices in energy efficiency market transformation evaluation (NMR, 2013). The second criterion is a set of evaluation quality standards prescribed by the Joint Committee on Standards for

Educational Evaluation *Program Evaluation Standards* (PrgES) (2010). These standards have wide usage in the evaluation discipline and have been adopted by both the American Evaluation Association and the Canadian Evaluation Society.

Market Transformation Evaluation Best Practices

A literature review to gain clarity on market transformation evaluation located a few useful resources. In particular, the work by NMR, of Somerville, MA, for the California investor-owned utilities provided an informed model of market transformation and its evaluation. Their reporting suggests a set of best practices that, on the surface, have face validity for being a comprehensive set of 10 practices that, if followed, would provide useful evaluation of MT programing (NMR, 2013). By accepting that the NMR reporting does a good job of describing what a good MT evaluation looks like, this provides a "standard" against which to review MRE MT evaluation to see if it is "a good rendition" of market transformation evaluation. There is some issue with selecting this model. The NMR research was developed for utilities that were engaged in attribution-related evaluation, and NEEA works in an ecosystem that bounds NEEA to claim only limited attribution to influencing the market. This shortcoming is accounted for in how the model is used herein to review MRE's MT evaluation.

Program Evaluation Standards

The PrgES provide an independent framework for rating the MRE evaluation work in general. This set of standards, though born and raised in the education evaluation domain, has been adopted and adapted in many other domains and in other parts of the world. Some of the standards are less applicable to this situation, but overall, there are useful insights to be gleaned by thinking through what MRE does against these standards. There are 5 standards with a varying number of statements attached to each standard so that 30 statements are provided to apply to an evaluation. The 5 standards are:

1. Utility
2. Propriety
3. Feasibility
4. Accuracy
5. Evaluation Accountability

Dan Stufflebeam was the founding chair of the Joint Committee on Standards for Educational Evaluation (JCSEE) and developed numerous tools to support the application of the JCSEE standards, including a recently uncovered unpublished metaevaluation checklist (dated 2016). The checklist includes a scoring rubric one applies during analysis. This 2016 checklist is used herein to review and rate MRE's MT evaluation work.

Description of the Method and Scope of Work

The metaevaluation design takes a mixed approach. First, reviewing the program documentation and doing interviews provides a holistic perspective, then the analysis turns to a more deductive analysis of the program by comparing it to existing standards.

Through these combined efforts, it is hoped to provide an informed and useful description of MRE MT evaluations and identify some ways where the evaluation approach may be improved.

The Tasks

The research activities described in the four subtasks below comprise the assessment of MRE's approach to the evaluation of market transformation programs. The tasks are reviewing NEEA MRE MT evaluation documentation, interviewing key informants from the MT evaluation program teams, comparing the MT evaluation programming to best practices and an evaluation quality checklist, and coalescing the learning from these sources of information to report on MRE MT evaluations, including strengths and weaknesses.

Table 1: Tasks in the Scope of Work

Sub-Tasks	Activities
1. Review identified NEEA Reports and other resources	Attend an informal, virtual "meet and greet" with NEEA's MRE team. In collaboration with the NEEA MRE Sr. Manager, develop a list of resources to review. Thoroughly review the resources. Draft a brief memo summarizing insights and themes.
2. Interview MRE's MT evaluation stakeholders	Develop list of interviews and interview guide. Conduct interviews. Analyze data. Prepare topline report of interview results.
3. Analyze and synthesize results	Synthesize interview and document analysis data and develop preliminary findings.
4. Report findings	Develop report outline. Deliver draft report for review by MRE team. Conduct synthesis workshop with MRE team. Resolve comments from the review and submit revised draft. Finalize report.

Method

There are two groups of data sources, an organizing framework, a set of best practices, and a checklist of evaluation quality used in this study. The data sources are program documentation and related reports, and interviews. Data obtained from these were organized using the working logic of evaluation. Finally, a set of MT evaluation best practices (NMR, 2013) is compared to what MRE is doing and the metaevaluation checklist [PrgES] is used to assess the quality of MRE MT evaluations.

Document Review

Document Selection Process

The MRE Sr. Manager[4] presented a set of NEEA MT programs that would be well-positioned in their life cycle to give a variety of views into the MRE MT evaluation work. These programs included: the Ductless Heat Pump (DHP), Heat Pump Water Heater (HPWH), and Luminaire Level Lighting Controls (LLLC) programs. For each program, the metaevaluator was provided a variety of documents, including logic models, market progress indicator lists (MPI), program stories, product and program plans, and market progress evaluation reports (MPER). The programs and associated documents are listed in the table below. During further discussions in bi-weekly calls, other documents were identified and provided. The metaevaluator also did research and turned up resources on the NEEA website and on the internet more broadly to review[5].

Table 2: Program Documents Reviewed

	DHP	HPWH	LLLC
Market Characterization	X	X	X
Market Baseline			X
Product Plan		X	X
Program Plan		X	X
MT Story		X	X
Logic Model	X	x2	X
MPIs	X	X	X
MPER	x4	x5	
Cost Benefit Model	X		
Statement of Work	LTMT Contract	MPER6	MPER1

Document Review Method

Each document was read closely and annotated. These annotations were regularly revisited throughout the evaluation process to refresh understandings and to identify themes across documents. Because of the relative limited breadth of documents, a coding schema was not designed, instead an in-vivo markup model was used that included gathering key phrases and terms as well as annotation of the documents.

[4] Amy Webb

[5] Other documents reviewed: NEEA Strategic plan 2020-2024; studies by other agencies located through literature searches; various internal MRE working documents, such as slide presentations, gap analysis templates, and evaluation and research plan template

Interviews

Interview Selection Process

The document review provided the necessary information to develop key questions that would guide the interviews and the metaevaluator drafted a protocol, on which the MRE Sr. Manager commented. Further, the MRE Sr. Manager provided a list of possible interviewees and the metaevaluator, applying a systems lens (Williams, 2019), selected those on MRE's immediate boundaries and whose relationship to MRE was through the work they did evaluating MT programming.

Table 3: Interviewee by Role in Relation to MRE/MT PE

Role	N	Description of Role
Market transformation manager	1	Manages the market transformation of a product
MRE scientist	5	Lead research and evaluation on Program Team
Program manager	1	Manages transformation programming
Director	2	Direct analytics and market transformation
Planning analyst and manager	2	Design market transformation modeling

Cross-Resource Synthesis

Following first pass reviews of the documents and the interview transcripts, a cross-resource synthesis was facilitated through the process of describing the MT evaluation's working logic and through analyzing the MT evaluation programming using the identified best practices and the PrgES checklist.

Describing MRE MT Evaluation

Fournier (Fournier, 1995; Fournier & Smith, 1993) discusses evaluation as having both a formal general logic (Scriven, 2007) and a working logic. The general logic was used above to refine the description of this research effort as more metaevaluation than assessment. The latter is used in the next section to describe the MRE MT program evaluations and to help understand claims developed by MRE.

Analysis Against the PrgES

The analysis for this report reviewed each of the 30 PrgES statements using a checklist designed by Dan Stufflebeam based on the 2010 3rd edition of the *Program Evaluation Standards*. The analysis provided an opportunity to review each of the interviews and codes and the notes on each document to substantiate the checklist's statements of quality.

Taking the checklist to its designed conclusion, the evaluation garners a score. That score is categorized as excellent, very good, good, fair, or poor. Stufflebeam (2016) says "There is no magic formula for setting cut scores" and that these were designed from his experience of many years in professional evaluation. Therefore, the research reported herein uses the

system he created and reports those ratings in the below section titled "Evaluative Application of the PrgES".

Refinement of Interpretations

Built into the metaevaluation interpretation moment was a workshop with the MRE Team. Because this metaevaluation is of their work, they had the first opportunity to see the report and to provide perspective to findings and recommendations. Their input informed the report.

Describing MRE Market Transformation Program Evaluation

This section uses the working logic framework developed by Fournier (1995), to describe MRE's MT evaluation. Evaluation working logic is useful because it highlights how evaluative claims are inherently tied to the problem being addressed and the phenomenon that was developed to address the problem.

Table 4: Working Logic Components

Problem	What is the outcome of interest, what are we solving? (e.g., market transformation)
Phenomenon	How are we getting to the solution, what has been put into place to affect change? (e.g., MT programming)
Question	Answerable about the problem, given the phenomenon? (e.g., is the market transforming?)
Claim	What will you claim in relation to this problem? Value/outcomes/impact (e.g., market is transforming seemingly on its own, or likely will if we terminate MT programming)

Problem – Transforming the Energy Efficiency Market

The *problem* being addressed by MRE's MT evaluation is the transformation of the market for whatever product they are supporting. This means measuring market transformation indicators (as outcome proxies), some having distant links (long causal chain) from efforts to indicator. That is, the movement of the market towards a self-driven evolution to being more energy efficient. MRE is not evaluating program effects, per se, but measuring indicators that provide evidence of some distant market influence against which the program efforts can be linked.

The *problem* is refined early by market characterizations, baseline reviews, and market test assessments. These set the stage for market progress evaluation reporting and later long-term monitoring and tracking. These efforts make it so that the *problem* of market transformation is fairly well understood before market development begins.

Phenomenon – Market Transformation Programming

The *phenomenon* for MRE's MT evaluation, the thing that is being evaluated by MRE, is market transformation programming. This is best reflected in Figure 1 below. The initiative

life cycle graphic shows the MT program's three major phases (concept development, program development, and market transformation). MRE provides information in all phases of the lifecycle. MRE's role at the Program Advancement milestone, the Program Development Check-In meeting, and other program lifecycle milestones is to ensure Market Progress and Evaluation Reports (MPERs) are designed, to as closely as possible, record the movement of the market in relation to indicators expected in advance. MPERs assess movement of barriers, and in some cases, effectiveness of activities, and the achievement of outcomes operationalized through MPIs.

The *phenomenon*, then, is the market transformation program and the evaluation of that is represented by market characterizations, baseline reviews (repeated), market test assessments, market progress and evaluation reports (repeated), and long-term monitoring and tracking (repeated). Each of these provides evaluative moments. For instance, the MPERs tell us if the MPIs have reached the expected (pre-defined in early stages of the ILC) standard at which point the Program transitions to an observation stance through the Transition to Long-Term Monitoring and Tracking Milestone, rather than an engagement stance.

Figure 1. Initiative Life Cycle Graphic

Questions – Is the market transforming as expected? Are we doing good work?

The *questions* section of the description of the MRE MT program evaluation efforts focuses on the undergirding of the MT program. The *questions* are defined by Market Transformation (MT) theory (which is used to define the phenomenon, or the program) and stem from the defined outcomes. In MT theory, outcomes are defined as "where the program wants the market to go." Programs use market progress indicators (MPIs) as what they expect to see if they are making progress towards some outcome, and metrics are the data they need to track progress on indicators towards an outcome.

The bottom-line *question* at the end of the program implementation stage is "has the market transformed into a more energy efficient market so that continuation will not provide a sufficient return to warrant the use of donor funds?" Other *questions* guide the MPERs on an irregular basis (e.g., outstanding research questions or risk assessment issues from the Product Plan), meanwhile the main *question* remains, "is the market moving in the direction we predicted in our program theory?"

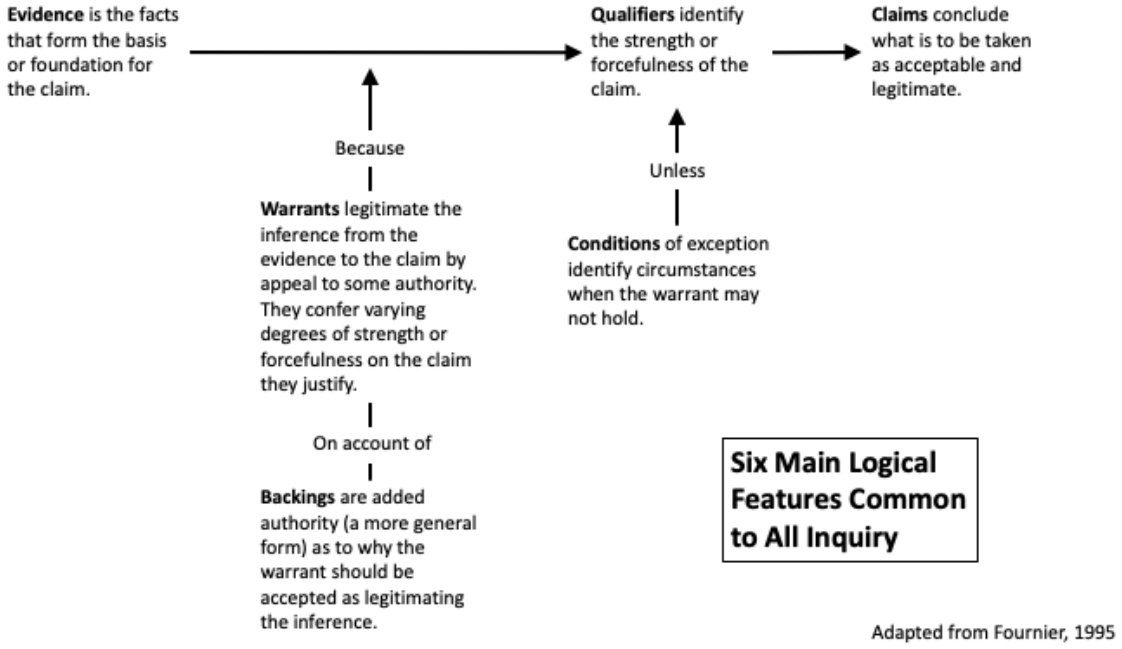
Claims – The market is evolving on its own. We are diligent in our efforts.

Claims are those statements that indicate taking some responsibility for a change. Deciding to move to long-term monitoring and tracking involves substantiating a *claim* that the product market looks to be on a trajectory to sustain its own forward momentum. On a single MPER, *claims* are made on a small scale, generally on the order of "MPIs are moving along as predicted" or "some parts of the market are not available to assess." Large scale *claims* come at decision points, such as the move from program implementation to long-term monitoring and tracking (e.g., "Findings from this MPER suggest that the market for DHPs in the Northwest will continue to transform without direct support from NEEA") or on even a grander scale as in NEEA's outward facing documents like its 2020-2024 Strategic and Business Plan ("…since 1996, the region has cost-effectively delivered more than 1,720 average Megawatts of efficiency through market transformation." (p.4)).

Each *claim* is supported by evidence and a warrant establishes why such evidence is sufficiently supportive of such a *claim*. Interrogating claims is the job of MRE Scientists. For example, when selecting an MPI the MRE Scientist critiques the premise that it is in fact good evidence of an outcome. One claim that should likely be interrogated on a regular basis is the idea that the MPIs satisfactorily constitute a "preponderance of evidence" as a proxy for indicating market progress towards transformation.

The below figure presents a visual representation of the general argumentation logic. Components of this logic are seen throughout MRE publications. They are explored in Appendix D to help draw out the various parts of evaluation *claims* and evidence.

Figure 2: Components of an Argument: Six Main Logical Features Common to All Inquiry



Contextual Factors – Definitions Differ

Argumentation claims may have different audiences and so evidence required (and warrants for such evidence) may also differ. Market transformation programming and its evaluation must keep in mind at least two, sometimes opposing, audiences. There is the internal decision-making apparatus that they are part of – the Program Team working to transform the market. More distant and external to that process is the accountability audience – the utilities and other NEEA funders, the Board of Directors, and others outside NEEA.

In the above examples, the evidence for delivering efficiency, while requiring several quantitative models with numerous assumptions, will witness intense scrutiny by utility partners and other vested stakeholders. NEEA's energy savings forecast cost-benefit model is regularly reviewed for each MT program. New purchases, sales, or installations in areas wherein a utility is providing an incentive are not claimed by NEEA. The evidence for claims of movement towards specific program outcomes, on the other hand, will be useful for informing programmatic decisions, and may be less scrutinized by external audiences.

Assessment of MRE MT Program Evaluation

Now that a thorough description has been developed of the MRE MT program evaluation efforts, an assessment can be advanced. This section provides an assessment of the MRE MT program evaluation efforts against both NMR's (2013) research-informed recommended practices and the premier standards in the evaluation discipline, the Joint Committee on Standards for Educational Evaluation *Program Evaluation Standards*

(PrgES). Reviewing against these two tools provides opportunity to further describe MRE MT program evaluation efforts. It provides a snapshot of how well the MRE evaluation efforts reflect the discipline's standards. Secondly, the process helps uncover perspectives on evaluation that can be useful when assessing the MRE MT approach. The completed metaevaluation checklist is provided in Appendix C and only tables listing the ratings applied appear in this section.

Comparing MRE to NMR Practice Recommendations

The MT programming *phenomenon* is developed in accordance with market transformation theory, which is reflected in the set of practices developed from the MT literature and listed in the NMR report (2013). Except for a few definitional differences, noted below, this set of practices is the most appropriate MT programming guidance available.

1. Match the evaluation strategy to the program logic – MRE Scientists are research and evaluation functional leads on the program team. They participate in the development of the program strategy, and they develop the MPIs. The MPIs are derived from the program logic. They also guide the MPERs. Participation in the program development through market characterization and other research efforts means the MRE Scientist is helping build a viable program that has some measurable indicators. They engage with the cross-functional Program Team, including Market Transformation Managers (MTMs), Market Analysts, Program Managers, and their own team of MRE Scientists through regular meetings. In these meetings, MRE Scientists continually ask evaluative questions, such as how one can measure certain expected outcomes, the relative value of getting data to answer a specific question and the cost to get that answer. MRE Scientists also address questions about the theory of change underpinning the efforts suggested to transform the market. For instance, how do you know if having more trained installers in a 50-mile radius is a good indicator for an improved market competition? A few products evolve from these early meetings, all focused on being more specific about the market and the effort towards its transformation: a logic model, Market Progress Indicators list (MPIs), a Product Plan (very useful description of the transformation plan), the MT Story, and the program plan (which is keyed to the logic model). Together, this documentation paints a picture of the program and provides the basis for MRE Scientists, in collaboration with the Program Team to build an evaluation plan. For older programs, documentation of that evaluation plan can be inferred in the MPERs. For more recent programs (and going forward for all) the evaluation plan is made explicit in the Market Research and Evaluation Plan. These MR&E Plans have only recently been given full attention and are being integrated across programs. This step is substantiated.

2. Track indicators tied to expected outcomes – The program MPIs operationalize the expected outcomes and these MPIs are regularly measured in the MPER. By operationalization, this means that, in the abstract, it is not obvious how every outcome will be measured and there may be numerous data elements, metrics, or measures, that define any particular outcome. For instance, a large outcome such as

"*there is strong market competition for providers of building commissioning*" requires some deeper definitions of its components (e.g., competition, strong market, providers) and the process of putting refined measure behind each of these is the process of operationalizing an outcome…and creating MPIs. Whether or not the identified MPIs are adequate to fully measure the outcome, whether there may be more leading instead of lagging indicators, and even whether these are *the right* indicators are just some of the many questions MRE Scientists are expected to instigate. This step is substantiated.

3. Perform regular, ongoing research into the status of the market – Annual MPERs provide timely insight into the market. MRE hires external vendors to do much of the research requested of them. This is a challenging model because not everyone understands MT theory as well as those at NEEA. This highlights a critical skill required of MRE Scientists, that of project manager (i.e., writing good RFPs to get what they need, managing relationships with vendors to ensure the questions are answered appropriately, and the politics of the space are constantly reviewed). This step is substantiated.

4. Assess market effects periodically – The NMR research, from which this framework was borrowed, uses the language "market effects", implying a rather direct cause-and-effect relationship between transformation programming and changes on indicators. Herein, the term "influence" is used because it is a more honest descriptor of what can be tracked and what can be claimed. Tracking MPIs from the outset is critical to ensuring NEEA keeps a pulse on the market indicators. This step is substantiated.

5. Refine the program theory and logic model – Annual MPERs provide the opportunity to review logic models and program theory. The MRE Scientists use these opportunities to not only measure MPIs, but also engage in further research, such as reviewing logic models or baseline data assumptions. This step is substantiated.

6. Assess attribution – NMR (2013) uses this definition of attribution developed by Rosenberg and Hoefgen (2009): "link program activities to identified market change in order to establish causality." (p. 27). Rosenberg and Hoefgen define the causal warrant for claiming cause, that is most like NEEA's approach to claiming influence: "assessing the consistency of the observed fact pattern with linkages predicted by the program logic model." (p. 78). While this causality linkage is not without merit, there is a lot of error that is likely unaccounted for in such a claim. Given the messy nature of markets and human behavior, NEEA chooses to be more reserved than claim attribution and prefers the term influence. They do not assess attribution, per se, rather they measure indicators and take responsibility for having influence in market movements directly related to their pre-defined indicators. It is this last point, "pre-defined indicators" that gives this claim strength: because NEEA puts forth expectations for movement on data points in advance, a claim of influence is

supported when that movement occurs after the programming is started. Within the parameters of this slightly modified framework, this step is substantiated.

7. Calculate net savings at the market level – the energy savings forecast cost-benefit model uses very conservative assumptions to be sure to not overestimate contribution. This step is substantiated.

8. Assess sustainability and prepare for exit or transition – Target saturation in the market and other indicators of "market uptake" are used to decide when to go into long-term monitoring and tracking. When the MPER shows satisfactory movement of the market for the investment, a decision is made that "the market is sufficiently transformed so that our continued investment will not generate a worthwhile return" and the program moves to an observational stance wherein long-term monitoring and tracking replaces MPERs. This step is substantiated.

9. Tell the market transformation story – there is a product story, but this is done early, and it is not clear how much this storytelling continues later in the life of the program. Telling the story of a market transformation can be very useful for others looking to replicate the effort, or for those unfamiliar with the work to get an informative picture of a transforming market and the key moments of that transformation. The recent efforts to require the Market Research and Evaluation Plan more broadly implies that this will be substantiated at some point in the near future.

10. Continue tracking market effects after the program has ended – Here again, the term effects would be replaced with influence, the tracking of which is the purpose of long-term monitoring and tracking. This step is substantiated by the recent move of DHP into LTMT.

Evaluative Application of the PrgES

As seen in Table 5, MRE MT evaluation scored very good to excellent, in all categories of quality. The evaluation work is highest on propriety and lowest on evaluation accountability. Averaging the ratings gives an overall score of 86%, or "very good" from Stufflebeam's perspective. Figure 3 provides a breakout of the statements within each standard.

Table 5: PrgES Ratings by Standard

Standard	Score
Propriety	96% - Excellent
Utility	91% - Excellent
Accuracy	88% - Very good
Feasibility	88% - Very Good
Evaluation Accountability	67% - Very Good
Overall	86% - Very good

Analyzing MRE MT evaluation programming against the PrgES distilled a few areas that could use some attention by the MRE department, and the report points out where a standard did not gain full marks. For an expanded view of the marks across the PrgES, please see the *Ratings* in Appendix C.

Propriety – Excellent

The *propriety* standard of human rights and respect (P3) is one place MRE can rather easily maximize a dimension score. No mention was found, in any of the documents, of any evaluation standards. The American Evaluation Association and the Canadian Evaluation Society have endorsed the PrgES. Though they were written for use in educational evaluation, they have broad applicability. The AEA also has a set of guiding principles for evaluators that is a useful tool. Making a transparent commitment to the industry's principles and standards would add to the credibility of the MRE division and its evaluation efforts.

Utility – Excellent

In the *utility* category, there is much attention put on the evaluator competencies to be able to produce useful evaluations. While the MRE Scientists are seen as competent by their colleagues and were rated excellent on this item, interviews turned up a few key competencies as critical to their perceived success.

- MRE Scientists are expected to lead from a position of knowledge of market transformation evaluation.
- MRE Scientists should be capable of critiquing and discussing MT evaluation and research methods with the uninitiated.
- MRE Scientists should have deep knowledge of sampling and sample sizes in relation to confidence intervals.

MRE Scientists need to be the knowledgeable researcher and evaluator on the Program Team and that includes the ability to, when necessary, build the cross-functional team capacity for understanding research and evaluation methods. They must be not only good methodologists, but they must also be good communicators who can engage differently skilled professionals in difficult discussions about evidence, claims, and confidence.

"Attention to stakeholders" (U2) is another consistent theme in the PrgES. It may be of less importance for NEEA, but it came up often while using the PrgES and perhaps NEEA should think about whether or not they are engaging with a wide enough representation of stakeholders. For instance, U2c: "Search out & invite input from groups or communities whose perspectives are typically excluded, especially stakeholders who might be hindered by the evaluation" was marked with a question because it was unclear that MRE reaches very broadly for stakeholder engagement. It was not clear whose voice is not being heard in their evaluations. It is clear from evaluation theory that engaging those usually not heard can be useful in helping to inform operationalization of outcomes, or even identifying the

right outcomes. There is useful literature on critical systems heuristics and participatory evaluation that could guide conversations about power, privilege, positionality, and voice that may improve the equity of MRE MT evaluations.

Accuracy – Very Good

The standard "Justified conclusions and decisions" (A1) adds the notion of strengthening conclusions by entertaining plausible alternative explanations. From the interviews, discussions with the MRE Scientists, and the resources and documentation provided, no evidence was seen that this is done. "Valid information" (A2), also raises the concern about threats to validity, which could be strengthened by entertaining plausible alternative explanations. "Explicit evaluation reasoning" (A7) also suggests that evaluations should make efforts to investigate alternative explanations for observed findings. MRE should consider adding procedures to strengthen exploration of alternative explanations for observed findings.

Feasibility – Very Good

Two statements in the project management standard (F1) point directly at good evaluation planning, and one was marked absent because the Research and Evaluation Plan (R&E) is still in a nascent stage. The R&E Plan has the potential to be a very valuable tool. Evaluators are in a unique position on many teams: they are often the memory holder, the person that has the institutional knowledge of the phenomenon being evaluated and therefore should step into the role of documentarian. This comes up later as a responsibility equally tied to keeping good documentation of the evaluation (E1).

"Practical procedures" (F2) surfaced the question as to how much data NEEA has and how much they are paying to have gathered anew. There are numerous mentions in the MPERs of using "NEEA data" for analyses and one interviewee noted a concern that sometimes 3rd party contractors may come to NEEA for data to complete their contract with NEEA's MRE.

Evaluation Accountability – Very Good

One area MRE scored relatively lower on is internal metaevaluation (E2). This could be an analysis issue in that some of what MRE Scientists do can be classified as internal metaevaluation. Certainly, the MRE team discusses the work they do and reviews RFPs and other documentation of the MT evaluation. That said, this analysis did not find evidence of such efforts.

Figure 3: PrgES Ratings by Metaevaluator

Criterion	Score & Rating	*Graph of Merit*				
		Poor (0-1)	**Fair** (2-3)	**Good** (4)	**Very Good** (5)	**Excellent** (6)
U1 Evaluator Credibility	6 (E)					
U2 Attention to Stakeholders	5 (VG)					
U3 Negotiated Purposes	6 (E)					
U4 Explicit Values	6 (E)					
U5 Relevant Information	6 (E)					
U6 Meaningful Processes and Products	5 (VG)					
U7 Timely & Appropriate Communicating & Reporting	5 (VG)					
U8 Concern for Consequences & Influence	6 (E)					
F1 Project Management	5 (VG)					
F2 Practical Procedures	5 (VG)					
F3 Contextual Viability	5 (VG)					
F4 Resource Use	6 (E)					
P1 Responsive & Inclusive Orientation	6 (E)					
P2 Formal Agreements	6 (E)					
P3 Human Rights & Respect	5 (VG)					
P4 Clarity & Fairness	6 (E)					
P5 Transparency & Disclosure	6 (E)					
P6 Conflicts of Interest	6 (E)					
P7 Fiscal Responsibility	6 (E)					
A1 Justified Conclusions & Decisions	5 (VG)					
A2 Valid Information	5 (VG)					
A3 Reliable Information	6 (E)					
A4 Explicit Program & Context Descriptions	6 (E)					
A5 Information Management	6 (E)					
A6 Sound Design & Analyses	6 (E)					
A7 Explicit Evaluation Reasoning	4 (G)					
A8 Communication & Reporting	6 (E)					
E1 Evaluation Documentation	6 (E)					
E2 Internal Metaevaluation	1 (P)					
E3 External Metaevaluation	6 (E)					

Limitations

Though this research attempted to triangulate data sources and analysis instruments, the best practices used herein have not been tested as a comprehensive, valid, and reliable way of doing the best evaluation of MT programming. Some assumptions are made about its applicability, and these have been noted. A second limitation is that the PrgES are, by their nature, reductive and narrow in scope. They take very complex concepts and reduce something like "utility" to 8 statements and 6 statements for each of those. In all, the PrgES checklist used herein provides 180 statements that are marked as present or absent, and these are then summed to create a rating of excellent, very good, good, fair, or poor. Many statements seemed to be applicable in only certain circumstances and still others may seem irrelevant to the MRE team. This report tries to acknowledge where standards were particularly extraneous or irrelevant. It takes a liberal stance on these and gives points where items were irrelevant and marked them with an asterisk [+*], adding notes where appropriate. The full ratings with notes are in Appendix C. Finally, it was realized in retrospect that those chosen for interviews (other than MRE team members) mostly experienced MRE through a small subset of the current MRE Scientists. For this reason, some of the recommendations and claims about MRE Scientist-needed skillsets are given as broad and soft. It seemed inappropriate during analysis to claim MRE Scientists had some shortcoming because others' assessment of capacity was driven by interactions with a small subset of MRE Scientists.

Summary and Conclusions

This section is organized by the three guiding metaevaluation questions put forth above.

Question 1: How do MRE MT evaluation practices compare to industry best practices as framed by NMR (2013)?

After reviewing MRE's MT evaluation against the best practices as published by NMR (2013), this report finds extremely laudatory practices. A great deal of effort is put into building, implementing, evaluating, and managing market transformation programs. There is regular reflection incorporated into the program, including the use of facilitation tools to develop and revise MT theory, so that data and questions are applied to keep focused on outcomes. MRE treats market transformation as a program that has an underlying logic around which evaluation questions are developed and answered. The theory of change embedded in the market progress indicator tables and the connection of the indicators to the logic model as operationalized outcomes makes transparent the linkages between actions and outcomes and defines expectations of the market. The program theory is established and revisited at regular intervals so that adjustments to MT efforts can be implemented. The key decision point of moving from program implementation to long-term monitoring and tracking is evaluated annually.

A strength of the MRE MT evaluation is that it reflects MT evaluation best practices (NMR, 2013). The constant engagement by MRE Scientists in the program development and the iterative sharing of data with the Program Team provides an environment where rigor and knowledge use are likely very high.

Because MRE Scientists hire 3rd party consultants to do the bulk of the research and evaluation work, **the skills required for this process (e.g., RFP writing, relationship management, contextual political awareness) should not be overlooked**.

Documenting and telling the market transformation story is a key component of evaluating MT phenomena. **MRE is making progress on this aspect by requiring the Market Research and Evaluation Plans be updated annually. This should be continued.**

Question 2: How do MRE MT evaluation practices score on the Joint Committee for Standards on Educational Evaluation Program Evaluation Standards?

MRE's MT evaluation scored *very good* when rated on the Joint Committee on Standards for Educational Evaluation *Program Evaluation Standards* (PrgES). MRE MT evaluation scored highest on propriety, followed by utility, accuracy, feasibility, and finally evaluation accountability. Though rated highest on propriety, no mention was found of MRE's commitment to evaluation standards and ethical principles. **Incorporating explicit statements about adherence to standards and ethical principles has the potential to improve the MRE work**. Stakeholder involvement is a strong theme in the PrgES and while the analysis in this research focused less on stakeholder engagement **MRE is encouraged to look beyond the current stakeholders and see whose voice is missing at the evaluation table** from the earliest phases of the program life cycle.

The accuracy standards indicated that **claims could be improved by entertaining plausible alternative explanations.** No evidence was seen that this is done. MRE Scientists can add this to their repertoire in the work with the Program team.

Evaluation accountability could be improved if the internal metaevaluation work of the MRE team is made more transparent and explicit. It is likely that the renewed focus on the Research and Evaluation Plan can bring an improved evaluation accountability.

Question 3: What recommendations can be developed from the comparison to best practices and the PrgES evaluation quality checklist?

The use of a learning model, where MRE Scientists are partners in ensuring decision-makers can learn more about their market and make changes when needed reflects the core of formative evaluation. It becomes summative at key decision moments, such as the decision to move to LTMT. Embodying this iterative learning model is critical in the complex and potentially volatile energy efficiency market. NEEA built a smart life cycle timeline that organizes the learning process and reflects the tenets of adaptive

management principles. Though not all are the purview of MRE, there are likely five key evaluative moments reflected in the program life cycle: opportunity advancement, concept advancement, program advancement, transition to LTMT, and monitoring complete. **Each of these are opportunities to implement evaluative thinking around choices, and NEEA is encouraged to consider how MRE can bring evaluative thinking to all these moments**. For instance, MRE could add procedures where MPERs are critically reviewed by those outside the Program Team, providing an opportunity to expand stakeholder voice, build capacity, and improve understanding.

Some improvement might be made around building out the story of the market transformation. MRE Scientists are central to the decision-making process that undergirds the program development and implementation. They embody the knowledge that goes into so many choices, and it should be their responsibility to document that understanding. The Research and Evaluation Plan template looks like a good start, but it seems that is only recently ascended to a universally used tool. This should change and perhaps **the MRE Scientist should take responsibility for using the R&E Plan as a storytelling device that documents the transformation of a market**. An example of a historical documentation that could be used to think about future evaluation documentation came up in an interview and is reflected in an HPWH MPER #5 recommendation.

- *With each MPER, NEEA should update the MPI tables with any new MPI values that may have been collected as part of the MPER. Maintaining up-to-date tables or lists of MPIs and their measurements will facilitate NEEA's efforts to track market progress in the future and to compare the results with past measurements. Having a clear history of the MPI values, how they were operationalized in various MPERs, and by what method they were gathered, will also facilitate MPER and market research planning (as many MPIs can be measured in conjunction with other market research activities, and do not necessarily need to be measured as part of an MPER).*

Telling the MT story in a way that helps define key moments where the trajectory may have shifted can shed light on how some efforts have more leverage than others. There is a rich literature on storytelling, and one way that has worked for programs is to build a character and tell the story of that character moving through the intervention. Around this character, you can provide data that may represent the larger market transformation, but it is grounded in the experience of the story's character.

No evidence was found of the MRE Team stating its guiding principles or standards. **The American Evaluation Association Guiding Principles and the PrgES used in this metaevaluation are useful tools that the MRE Team should bring into regular use** in both their work in developing MT evaluation and their outsourcing of the MPERs.

Finally, regarding the potential for methodological improvements to MRE's MT evaluation programming, we will never get 100% confidence in measuring things like market

transformation or even indicators of a single market transformation outcome. **Contribution analysis[6],[7] could be used to strengthen the credibility of how the preponderance of evidence is used to tell the MT programming story.**

If they are diligent to think through all the ways they may be wrong about claims, gather divergent voices in critiquing their assumptions, then perhaps they can develop a modest claim, through a preponderance of evidence identified in advance, and be comfortable in that claim. MRE is doing a fine job of being at the table and incorporating evaluative thinking at important moments. Proactive application of communication, negotiating, and methodological skills will be key to providing useful service in the "rough ground" of frontline market transformation evaluation practice.

[6] https://www.betterevaluation.org/en/plan/approach/contribution_analysis
[7] https://www.evaluationinnovation.org/wp-content/uploads/2021/03/Contribution-Analysis-Design_LauraHopkinsITAD.pdf

Appendix

A. References

Fournier, D. M. (1995). Establishing evaluative conclusions: A distinction between general and working logic. *New Directions for Evaluation, 1995*(68), 15–32. https://doi.org/10.1002/ev.1017

Fournier, D. M. and Smith, N. L. (1993). Clarifying the merits of argument in evaluation practice. *Evaluation and Program Planning* (16), 315–323.

Gullickson, A. M. (2020). The whole elephant: Defining evaluation. *Evaluation and Program Planning* (79),

Nevius, M., Hoefgen, L., Wilson-Wright, L. and Browne, C. (2013). *A Review of effective practices for the planning, design, implementation, and evaluation of market transformation efforts.* NMR Group, Inc. www.calmac.org https://www.calmac.org/publications/FINAL_NMR_MT_Practices_Report_20131125.pdf

Nevius, M., Hoefgen, L., Wilson-Wright, L. and Smith, B. A., Samiullah, S., Nguyen, L., and Prahl, R. (2015). *Effective practices for the evaluation of market transformation efforts.* International Energy Program Evaluation Conference, Long Beach, CA.

NMR Group, Inc. (2019). *Action plan for measuring market effects*. https://www.nmrgroupinc.com/action-plan-for-measuring-market-effects/

Northwest Energy Efficiency Alliance (2020). *NEEA Strategic and Business Plans 2020-2024*. Downloaded from http://www.neea.org/ August 11, 2021.

Scriven, M. (2007). The logic of evaluation. In H.V. Hansen, et. al. (Eds), *Dissensus and the Search for Common Ground*, CD-ROM (pp. 1-16). Windsor, ON: OSSA.

Stake, R. E. (1977). The countenance of educational evaluation. In A. A. Bellack, & H. M. Kliebard (Eds.). *Curriculum and evaluation* (pp. 372–390). Berkeley, CA: McCutchan.

State and Local Energy Efficiency Action Network. (2012). *Energy Efficiency Program Impact Evaluation Guide.* Prepared by Steven R. Schiller, Schiller Consulting, inc., www.seeaction.energy.gov

Stufflebeam, D. L. (2016). *Program Evaluations Checklist*. Unpublished manuscript.

Williams, B. (2019). *Systemic evaluation design: A workbook 2nd Ed.*

Yarbrough, D.B., Shula, L.M., Hopson, R.K., & Caruthers, F.A. (2010). *The Program Evaluation Standards: A guide for evaluators and evaluation users* (3rd. ed). Thousand Oaks, CA: Corwin Press.

B. Summary of Qualifications for Michael Harnar, PhD

Michael Harnar, Ph.D., is the owner of Pointed Arrows Consulting and the Interim Director of the Interdisciplinary PhD in Evaluation program at Western Michigan University. He obtained his doctorate in Psychology with an emphasis in Evaluation and Applied Research Methods at Claremont Graduate University and has worked on evaluation projects as diverse as educational programs for underserved minorities, community college program review and planning for accreditation, health-related website evaluations, railroad safety programs, community-focused non-profit capacity building grants, program and curriculum development for an online Master of Arts degree, and youth participatory evaluations. Dr. Harnar studies how humans interact with one another and with the world in which we live, especially when it comes to evaluating. His current research interests include studying how evaluators assure quality in their work, improving how evaluators are educated at the graduate level, developing empirical knowledge about metaevaluation practice, and describing how evaluation is used by commissioners. He is an evaluation consultant with more than 18 years of experience supporting complex initiatives with an emphasis on visioning, outcomes definition, and measurement. In 2011, an evaluation Dr. Harnar co-authored for the Federal Railroad Administration won the American Evaluation Association's Best Evaluation award. As an assistant professor, he teaches evaluation-focused courses in an evaluation interdisciplinary doctoral program and advises students on their research and evaluation dissertations. He founded Pointed Arrows Consulting, as a Veteran Owned Small Business, to support his independent evaluation work. One of his current research projects is a systematic review of metaevaluation practice where he is working to develop a descriptive theory of metaevaluation practice. Before entering the evaluation field, he spent more than 20 years in the television field, including serving as a Navy combat photographer and videographer and a freelance video cameraman in New York City.

As a practicing evaluator, Dr. Harnar's values are exemplified by the guiding principles authored by the American Evaluation Association:

- Systematic inquiry: conduct data-based inquiries that are thorough, methodical, and contextually relevant.
- Competence: provide skilled professional services to the project stakeholders.
- Integrity/honesty: behave with honesty and transparency to ensure the integrity of the evaluation.
- Respect for people: honor the dignity, well-being, and self-worth of individuals and acknowledge the influence of culture within and across groups.
- Common Good and Equity: strive to contribute to the common good and advancement of an equitable and just society.

This report is the culmination of work produced entirely by Dr. Harnar. Where others were consulted or they provided input on findings, that is made clear in the writing to recognize the contribution.

C. Detailed Review of MRE MT Evaluation Against the PrgES

Northwest Energy Efficiency Alliance
Market Research & Evaluation
Market Transformation Evaluation Ratings

Systematic rating by against the 30 Joint Committee (2010) *Program Evaluation Standards.* Judged the NEEA MRE MT evaluation against each standard by marking each checkpoint plus (+) if met, minus (-) if not met, and question mark (?) if insufficient information was available to make a judgment. To enhance consistency of judging, if a checkpoint is not applicable, assigned (+*). Notes are provided herein for each (-) and where necessary (+*).

Utility
*THE **UTILITY STANDARDS** ARE INTENDED TO ENSURE THAT AN EVALUATION IS ALIGNED WITH STAKEHOLDERS' NEEDS SUCH THAT PROCESS USES, FINDINGS USES, AND OTHER APPROPRIATE INFLUENCES ARE POSSIBLE.*
U1 Evaluator Credibility. [*Evaluations should be conducted by qualified people who establish and maintain credibility in the evaluation context.*]
a. [+] Engage evaluators who possess the needed knowledge, skills, experience, and professional credentials b. [+] Engage evaluators whose evaluation qualifications, communication skills, and methodological approach are a good fit to the stakeholders' situation and needs c. [+*] Engage evaluators who are appropriately sensitive and responsive to issues of gender, socioeconomic status, race, language, and culture d. [+] Engage evaluators who build good working relationships, and listen, observe, clarify, and attend appropriately to stakeholders' criticisms and suggestions e. [+] Engage evaluators who have a record of keeping evaluations moving forward while effectively addressing evaluation users' information needs f. [+] Give stakeholders information on the evaluation plan's technical quality and practicality, e.g., as assessed by an independent evaluation expert
[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor
Comments re. U1, as appropriate: a. Some competencies are highlighted by interviewees, such as the ability to moderate challenging conversations that ask colleagues to question their assumptions when defining MPIs and market transformation more generally. Competency worth also noting: knowing when to get more specific and where

to move quickly because the market may be changing quickly. Also, negotiation and relationship management, can't get enough of the latter 2.

b. The MRE Scientists all seem to be very interested in this arena and doing good work around energy efficiency.
c. What really is "appropriately sensitive and responsive"? I have little insight into this because it was not a line of inquiry, nor did it come up as a topic during interviews. Except for the culture of energy efficiency. This was well-attended to.
d. The MRE each talked about relationships and the importance of working well with, while appropriately challenging the Program team.
e. While this did not come up as a consistent issue, there was some discussion about the critical importance of this skill. MRE Scientists need to be practiced at the business of contract and project management.
f. There are some limits to how much "limitations" are comfortably included in MPERs and MRE reporting more generally. Concerns over key constituency perception of "turf" seems relevant. Some talk about the extreme underestimating of contribution by giving so much credit to the utility that is providing a rebate.

U2 Attention to Stakeholders. [*Evaluations should devote attention to the full range of individuals and groups invested in the program or affected by the evaluation*.]

a. [+] Clearly identify and arrange for ongoing interaction with the evaluation client
b. [+*] Identify and arrange for appropriate exchange with the other right-to-know audiences, including, among others, the program's authority figures, implementers, beneficiaries, and funders
c. [?] Search out & invite input from groups or communities whose perspectives are typically excluded, especially stakeholders who might be hindered by the evaluation
d. [+] Help stakeholders understand the evaluation's boundaries and purposes and engage them to uncover assumptions, interests, values, behaviors, and concerns regarding the program
e. [+] Determine how stakeholders intend to use the evaluation's findings
f. [+] Involve and inform stakeholders about the evaluation's progress and findings throughout the process, as appropriate

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. U2, as appropriate:

a. MRE has a range of stakeholders, not only the "client", which may be considered the Program team. There are also utilities following and vested in the work of MRE and so the reporting must take into account not just the Program team. A distant, but important stakeholder could be the utility customers that ultimately fund the NEEA work. NEEA has chosen to be invisible to consumers. This deliberate choice certainly has roots; which I have not uncovered. Are market actors involved in deciding market indicators and MPIs to be used as proxies for market transformation?

b. What are the "right-to-know" audiences for MRE evaluations? This seems like it could be regulators and others that may have a more indistinct interest in NEEA's work.

c. This may be a part of the process; it was not a line of inquiry for me. I'll be curious to hear what the MRE Scientists think about this item.

d. This is a competency that is much needed in this work – to "help stakeholders understand the evaluation's boundaries". MRE Scientists are regularly involved in these conversations and Amy has given presentations on the life cycle of the MRE work to expand knowledge within NEEA about the work they do. This may be done to a lesser extent with external stakeholders.

e. This seems embedded in the process. In fact, the very nature of the process, that of learning to inform program planning, is intended to inform use. Here is a bit of the adaptive management perspective – use for learning is primary for most of the work.

f. This certainly happens with the Program team, perhaps less so with more external stakeholders.

U3 Negotiated Purposes. [*Evaluation purposes should be identified and revisited based on the needs of stakeholders.*]

a. [+] Identify the client's stated purposes for the evaluation
b. [+] Engage the client and stakeholders to weigh stated evaluation purposes—e.g., against their perceptions of dilemmas, quandaries, and desired evaluation outcomes—and to embrace evaluation's bottom-line goal of assessing value, e.g., a program's merit, worth, or significance
c. [+] Help the client group consider possible alternative evaluation purposes, e.g., program planning, development, management, and improvement; program documentation and accountability; and judging the program's quality, impacts, and worth

d. [+] Engage the client to clarify and prioritize the evaluation's purposes using appropriate tools such as needs assessments and logic models
e. [+] Provide for engaging the client group periodically to revisit and, as appropriate, update the evaluation's purposes
f. [+] Assure that initial and updated evaluation purposes are communicated to the full range of stakeholders

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. U3, as appropriate:

The ongoing, early, and engaged discussions with Program team; development of MPIs, Logic Model; all inform the purposes. It seems that attending to the outcome of readiness for long-term monitoring and tracking is agreed upon by the Program team.

U4 Explicit Values. [*Evaluations should clarify and specify the individual and cultural values underpinning the evaluation purposes, processes, and judgments.*]

a. [+] Make clear the evaluator's commitment to certain, relevant values, e.g., an evaluation's utility, feasibility, propriety, accuracy, and accountability and a program's equity, fairness, excellence, effectiveness, safety, efficiency, fiscal accountability, legality, and freedom from fraud, waste, and abuse
b. [+] Engage the client and program stakeholders in an effective process of values clarification, which may include examining the needs of targeted program beneficiaries, the basis for program goals, and the rationale for defined evaluation purposes
c. [+] Assist the client group to air and discuss their common and discrepant views of what values and purposes should guide the program evaluation
d. [+] Acknowledge and show respect for stakeholders' possibly diverse perspectives on value matters, e.g., by assisting them to seek consensus or at least reach an accommodation regarding possible alternative interpretations of findings against different values
e. [+] Clarify the values that will undergird the evaluation, taking account of client, stakeholder, and evaluator positions on this matter
f. [+] Act to ensure that the client and full range of stakeholders understand and respect the values that will guide the collection, analysis, and interpretation of the evaluation's information

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. U4, as appropriate:

NEEA is committed to diversity, equity, and inclusion. MRE MT evaluations are committed to being useful, accurate, proper, feasible, and accountability. The last is evidenced by this metaevaluation/assessment of their work. Program values are embedded in the logic model and MPIs.

U5 Relevant Information. [*Evaluation information should serve the identified and emergent needs of intended users.*]

a. [+] Interview stakeholders to determine their different perspectives, information needs, and views of what constitutes credible, acceptable information
b. [+] Plan to obtain sufficient information to address the client group's most important information needs
c. [+] Assess and adapt the information collection plan to assure adequate scope for assessing the program's value, e.g., its merit, worth, or significance
d. [+] Assure that the obtained information will address and keep within the boundaries of the evaluation's stated purposes and key questions
e. [+] Allocate time and resources to collecting different parts of the needed information in consideration of their differential importance
f. [+] Allow flexibility during the evaluation process for revising the information collection plan pursuant to emergence of new, legitimate information needs

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. U5, as appropriate:

a. The MPIs are generally accepted and addressed by the MPERs. The information obtained through the MPERs are directly from the MPIs, making it quite relevant.
b.
c.
d.
e.
f. There is some criticism of MRE's flexibility to adapt to the evolving nature of market transformation.

U6 Meaningful Processes and Products. [*Evaluation activities, descriptions, findings, and judgments should encourage use.*]

a. [+] Budget evaluation time and resources to allow for meaningful exchange with stakeholders throughout the evaluation process
b. [-] Engage the full range of stakeholders to assess the original evaluation plan's meaningfulness for their intended uses
c. [+] During the evaluation process, regularly visit with stakeholders' to assess their evaluation needs and expectations, also, as appropriate, to obtain their assistance in executing the evaluation plan
d. [+] Regularly obtain stakeholders' reactions to the meaningfulness of evaluation procedures and processes
e. [+] Invite stakeholders to react to and discuss the accuracy, clarity, and meaningfulness of evaluation reports
f. [+] As appropriate, adapt evaluation procedures, processes, and reports to assure that they meaningfully address stakeholder needs

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. U6, as appropriate:

a. There is considerable interchange between the internal Program team stakeholders.
b. This is questionable as to whether or not external stakeholders are engaged in the MPER development, writing, and publishing process. Outreach and engagement of a broader base of stakeholders could improve processes and interpretations.
c. See a. above.
d. See a. above.
e. MPERs are shared publicly. The results are shared and discussed with the Program team.
f. Done.

U7 Timeliness and Appropriate Communication and Reporting. [*Evaluations should attend in a timely and ongoing way to the reporting and dissemination needs of stakeholders.*]

a. [+] Plan to deliver evaluation feedback pursuant to the client group's projection of when they will need reports, but allow flexibility for responding to changes in the program's timeline and needs
b. [?] Plan, as appropriate, to give stakeholders access to important information as it emerges

c. [+] Employ reporting formats and media that accommodate the characteristics and serve the needs of the different audiences
d. [+] Determine how much technical detail to report by identifying and taking account of the audience's technical background and expectations
e. [+] Plan and budget evaluation follow-up activities so that the evaluator can assist the client group to interpret and make effective use of the final evaluation report
f. [+] Pursuant to the above checkpoints, formalize expectations for communicating and reporting to the sponsor and stakeholders in the evaluation contract

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. U7, as appropriate:

a. This seems to be a challenge in MT evaluation, and it is one criticism that MRE may hear from others. The "flexibility for responding to changes" in markets makes it difficult for programs to easily pivot, and for MRE to do cost-efficient evaluation of Programs.
b. I did not get the sense that MRE shared data as it emerged, rather in the MPER reporting as it is finalized.
c. MPERs do not come in formats other than reports.
d. Part of the job of the MRE Scientist working with the vendor evaluator.
e. Part of the functional lead of the MRE Scientist on the Program team.
f. Inherent.

U8 Concern for Consequences and Influence. [*Evaluations should promote responsible and adaptive use while guarding against unintended negative consequences and misuse.*]

a. [+] Identify the stakeholders' formal and informal communication mechanisms that connect stakeholders and, as appropriate, channel evaluation findings through these mechanisms
b. [+] Be vigilant and proactive in identifying and appropriately communicating with stakeholders who appear to be sabotaging the evaluation and, as necessary, counteract the sabotage
c. [+] Plan to meet, as appropriate, with stakeholders to help them apply findings in ways that are logical, meaningful, ethical, effective, and transparent
d. [+] In discussing evaluation findings with the client group stress the importance of applying the findings in accordance with the evaluation's negotiated purposes
e. [+] Be vigilant to identify, prevent, or appropriately address any misuses of

evaluation findings

f. [+] Follow up evaluation reports to determine if and how stakeholders applied the findings

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. U8:

None

Scoring the Evaluation for UTILITY **Add the following:**		**Strength of the evaluation's provisions for UTILITY:**
Number of Excellent ratings (0-8)	5 x 4 = 20	**[X] 29 (91%) to 32: Excellent**
Number of Very Good (0-8)	3 x 3 = 9	[ ] 21 (66%) to 28: Very Good
Number of Good (0-8)	0 x 2 = 0	[ ] 13 (41%) to 20: Good
Number of Fair (0-8)	0 x 1 = 0	[ ] 5 (16%) to 12: Fair
Total score:	= 29	[ ] 0 (0%) to 4: Poor
		29 ÷ 32 = .906 x 100 = 91%

Feasibility

*THE **FEASIBILITY STANDARDS** ARE INTENDED TO ENSURE THAT AN EVALUATION IS VIABLE, REALISTIC, CONTEXTUALLY SENSITIVE, RESPONSIVE, PRUDENT, DIPLOMATIC, POLITICALLY VIABLE, EFFICIENT, AND COST EFFECTIVE.*

F1 Project Management. [*Evaluations should use effective project management strategies.*]

a. [+] Ground management of the evaluation in knowledge of the stakeholders' environment and needs and the evaluation's purposes
b. [+] Prepare a formal management plan including, e.g., the evaluation's goals, procedures, assignments, communication, reporting, schedule, budget, monitoring arrangements, risk management arrangements, and accounting procedures
c. [+*] Recruit evaluation staff members who collectively have knowledge, skills, and experience required to execute, explain, monitor, and maintain rigor, viability, and credibility in the evaluation process
d. [+] Involve and regularly inform an appropriate range of stakeholders
e. [+] Systematically oversee and document the evaluation's activities and expenditures
f. [-] Periodically review the evaluation's progress and, as appropriate, update the evaluation plan and procedures

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. F1, as appropriate:

a. The MRE Scientists and the vendors they hire are familiar with the rough ground of energy efficiency and its market.
b. This seems embedded in the Market Research and Evaluation Plan, but this document does not have wide-spread use.
c. Certainly a skill worth noting and highlighting.
d. MRE serves the cross-functional Program Team.
e. MRE has systems in place to monitor vendors on contract.
f. While MRE does periodically review the evaluation process.

F2 Practical Procedures. [*The procedures should be practical and responsive to the way the program operates.*]

a. [+] Assess and confirm the program's evaluability before deciding to proceed with the evaluation
b. [+] Employ procedures that fit well within the program and its environment
c. [+*] Assure that the selected procedures take account of and equitably accommodate the characteristics and needs of diverse stakeholders
d. [+] Obtain relevant insider knowledge and incorporate it into the data collection process
e. [?] Make efficient use of existing information and avoid needless duplication in collecting data
f. [+] Conduct the evaluation so as to minimize disruption to the program

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. F2, as appropriate:

a. This is built into the MRE process as part of the Program Team.
b. Ditto
c. The MRE Scientists spoke of ensuring the data collected and reported on takes into account the utility partners and their perception of roles, such as who gets to claim subsidies, what is rural, and which states have agreements that may impinge upon data collection.
d. Vendor selection includes this sort of requisites.
e. There may be opportunities to capitalize off other data already in the NEEA data warehouse.

f. Managing the data collection to not disrupt the market, but to allow the processes to evolve while still collecting the information.

F3 Contextual Viability. [*Evaluations should recognize, monitor, and balance the cultural and political interests and needs of individuals and groups.*]

a. [+] Investigate the program's cultural, political, and economic contexts by reviewing such items as the program's funding proposal, budget documents, organizational charts, reports, and news media accounts and by interviewing such stakeholders as the program's funder, policy board members, director, staff, recipients, and area residents
b. [+] Take into account the interests and needs of stakeholders in the process of designing, contracting for, and staffing the evaluation
c. [+*] Enlist stakeholder and interest group support through such means as regular exchange with a review panel composed of a representative group of stakeholders
d. [+] Practice even-handedness and responsiveness in relating to all stakeholders, e.g., in the composition of focus groups
e. [+] Avert or identify and counteract attempts to bias or misapply the findings
f. [+*] Provide appropriate mechanisms for stakeholders to remain informed about the evaluation's progress and findings, such as an evaluation project website, an evaluation newsletter, targeted reports, and a telephone response line

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. F3, as appropriate:

None

F4 Resource Use. [*Evaluations should use resources effectively and efficiently.*]

a. [+] Negotiate a budget--ensuring that the contracted evaluation work can be completed efficiently and effectively—to include the needed funds and the necessary in-kind support and cooperation of program personnel
b. [+] Balance effectiveness and efficiency in resource use to help ensure that

the evaluation will be worth its costs and that sponsors will get their money's worth

c. [+] Use resources carefully with as little waste as possible
d. [+] Utilize existing data, systems, and services when they are well aligned with the evaluation's purposes
e. [+*] Document the evaluation's costs, including time, human resources, expenditures, infrastructure support, and foregone opportunities
f. [+*] Document the evaluation's benefits, including contributions to program improvement, future funding, better informed stakeholders, and dissemination of effective services

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. F4, as appropriate:

a.
b.
c.
d.
e. I did not gather evidence on this point.
f. This seems problematic for me. I don't think it is necessarily "efficient" use of resources to document the evaluation's benefits. It could be useful in EA1 evaluation documentation. But, I'm dubious about its value as good use of resources.

Scoring the Evaluation for FEASIBILITY Add the following:		**Strength of the evaluation's provisions for FEASIBILITY:**	
Number of Excellent ratings (0-4)	2 x 4 = 8	[ ] 15 (94%) to 16:	Excellent
Number of Very Good (0-4)	2 x 3 = 6	**[X] 11 (69%) to 14:**	**Very Good**
Number of Good (0-4)	0 x 2 = 0	[ ] 7 (44%) to 10:	Good
Number of Fair (0-4)	0 x 1 = 0	[ ] 3 (19%) to 6:	Fair
Total score:	**= 14**	[ ] 0 (0%) to 2:	Poor
		14 ÷ 16 = .875 x 100 = 88%	

Propriety

*THE **PROPRIETY STANDARDS** ARE INTENDED TO ENSURE THAT AN EVALUATION WILL BE CONDUCTED PROPERLY, FAIRLY, LEGALLY, ETHICALLY, AND JUSTLY WITH RESPECT TO (1) EVALUATORS' AND STAKEHOLDERS' ETHICAL RIGHTS, RESPONSIBBILITIES, AND DUTIES; (2) SYSTEMS OF RELEVANT LAWS, REGULATIONS, AND RULES; AND (3) ROLES AND DUTIES OF PROFESSIONAL EVALUATORS.*

P1 Responsive and Inclusive Orientation. [*Evaluations should be responsive to stakeholders and their communities.*]

a. [+] Acquire and take account of knowledge of the program environment's history, significant events, culture, and other factors affecting the program and its evaluation
b. [+*] Identify stakeholders broadly, gather useful information from them, and include them, as appropriate, in decisions about the evaluation's purposes, questions, and design
c. [+] Engage and serve the full range of stakeholders in an even-handed manner, regardless of their politics, personal characteristics, status, or power
d. [+] Design and schedule the evaluation to provide multiple opportunities for stakeholders to be involved, contribute, and be heard throughout the evaluation process
e. [+] Be open to and thoughtfully consider stakeholders' contradictory views, interests, and beliefs regarding the program's prior history, goals, status, achievements, and significance
f.[+*] Avert or counteract moves by powerful stakeholders to dominate in determining evaluation purposes, questions, and procedures and interpreting outcomes

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. P1, as appropriate:

None

P2 Formal Agreements. [*Evaluation agreements should be negotiated to make obligations explicit and take into account the needs, expectations, and cultural contexts of clients and other stakeholders.*]

[+] Negotiate evaluation-related obligations, with the client, including what is to be done, how, by whom, when, and at what cost

[+] Make ethical, legal, and professional stipulations and obligations explicit and binding regarding such evaluation matters as evaluation purposes and questions, confidentiality/anonymity of data, editorial authority, release of reports, evaluation follow-up activities, cooperation of program staff, funds and in-kind resources, and provision for a metaevaluation

[+] Employ the contract negotiation process to strengthen trust in communications through stakeholder consultation and, unless restricted by laws or regulations, allowing stakeholders to review the printed agreement

[+] Ensure that formal evaluation agreements conform to federal, tribal, state, or local requirements, statutes, and regulations

[+] Employ negotiated agreements to monitor, track, and assure effective implementation of specific duties and responsibilities

[+] Revisit evaluation agreements over time and negotiate revisions as appropriate

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. P2, as appropriate:

None

P3 Human Rights and Respect. [*Evaluations should be designed and conducted to protect human and legal rights and maintain the dignity of participants and other stakeholders.*]

a. [+] Adhere to applicable federal, state, local, and tribal regulations and requirements, including those of Institutional Review Boards, local/tribal constituencies, and ethics committees that authorize consent for conduct of research and evaluation studies
b. [+] Take the initiative to learn, understand, and respect stakeholders' cultural and social backgrounds, local mores, and institutional protocols
c. [-] Make clear to the client and stakeholders the evaluator's ethical principles and codes of professional conduct, including the standards of the Joint Committee on Standards for Educational Evaluation
d. [+] Institute and observe rules, protocols, and procedures to ensure that all evaluation team members will develop rapport with and consistently manifest respect for stakeholders and protect their rights
e. [+] Make stakeholders aware of their rights to participate, withdraw, or challenge decisions that are being made at any time during the evaluation process
f. [+*] Monitor the interactions of evaluation team members and stakeholders and act as appropriate to ensure continuing, functional, and respectful communication and interpersonal contacts throughout the evaluation

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. P3, as appropriate:

None

P4 Clarity and Fairness. [*Evaluations should be understandable and fair in*

addressing stakeholder needs and purposes.]

a. [+*] Develop and communicate rules that assure fairness and transparency in deciding how best to allocate available evaluation resources to address the possible competing needs of different evaluation stakeholders
b. [+] Assure that the evaluation's purposes, questions, procedures, and findings are transparent and accessible by all right-to-know audiences
c. [+] Communicate to all stakeholders the evaluation's purposes, questions, and procedures and their underlying rationale
d. [+] Make clear and justify any differential valuing of any stakeholders' evaluation needs over those of others
e. [+] Carefully monitor and communicate to all right-to-know audiences the evaluation's progress and findings and do so throughout all phases of the evaluation
f. [+] Scrupulously avoid and prevent any evaluation-related action that is unfair to anyone

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. P4, as appropriate:

a. No rules exist, but the decision about who to prioritize is regularly discussed.
b.
c.
d. There is some differential valuing going on between those developing the program and considering external utility and other audiences. The latter seem to take priority because of their connection to the funding.
e.
f.

P5 Transparency and Disclosure. [*Evaluations should provide complete descriptions of findings, limitations, and conclusions to all stakeholders unless doing so would violate legal or propriety obligations.*]

a. [+] Identify and disclose to all stakeholders the legal and contractual constraints under which the evaluation's information can be released and disseminated
b. [+] Maintain open lines of communication with and be accessible to, at least representatives of, the full range of stakeholders throughout the evaluation, so they can obtain the information which they are authorized to review
c. [+] Before releasing the evaluation's findings, inform each intended recipient of the evaluation's policies—regarding such matters as right-to-know audiences, human rights, confidentiality, and privacy—and, as appropriate, acquire their written agreement to comply with these policies
d. [+] Provide all stakeholders access to a full description and assessment of the program, e.g., its targeted and actual beneficiaries; its aims, structure, staff, process, and costs; and its strengths, weaknesses, and side effects
e. [+] Provide all stakeholders with information on the evaluation's conclusions and limitations
f. [+*] Provide all right-to-know audiences with access to information on the evaluation's sources of monitory and in-kind support

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. P5, as appropriate:

a. Assume standard MRE practice
b. Assume standard MRE practice
c. Assume standard MRE practice
d. Reporting of limitations of MPERs looks to be a weak spot.
e. Good
f. I did not inquire into this, but I could imagine this being challenging.

P6 Conflicts of Interests. [*Evaluators should openly and honestly identify and address real or perceived conflicts of interests that may compromise the evaluation.*]

a. [+*] Throughout the evaluation process search for potential, suspected, or actual conflicts of interest
b. [+*] Search for conflicts involving a wide range of persons and groups, e.g., those associated with the client, the program's financial sponsor, program recipients, area residents, the evaluator, and other stakeholders

c. [+*] Search for various kinds of conflicting interests, including prospects for financial gains or losses, competing program goals, alternative program procedures, alternative evaluation approaches, and alternative bases for interpreting findings
d. [+*] Take appropriate steps to manage identified conflicts so that the evaluation maintains integrity and high quality
e. [+*] Attend to conflicts of interest through effective communication with the client and other pertinent parties and in a spirit of mutual and deliberate understanding and learning
f. [+*] Document and report identified conflicts of interest, how they were addressed, and how they affected the evaluation's soundness

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. P6, as appropriate:

I did not inquire into this, but from conversations with Amy, this is likely part of the MRE process management.

P7 Fiscal Responsibility. [*Evaluations should account for all expended resources and comply with sound fiscal procedures and processes.*]

a. [+*] Plan and obtain approval of the evaluation budget before beginning evaluation implementation
b. [+*] Be frugal in expending evaluation resources
c. [+*] Employ professionally accepted accounting and auditing practices
d. [+*] Maintain accurate and clear fiscal records detailing exact expenditures, including adequate personnel records concerning job allocations and time spent on the job
e. [+*] Make accounting records and audit reports available for oversight purposes and inspection by stakeholders
f. [+*] Plan for and obtain appropriate approval for needed budgetary modifications over time or because of unexpected problems

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. P7, as appropriate:

I did not inquire into this, but from conversations with Amy, this is likely part of the MRE process management.

Scoring the Evaluation for PROPRIETY Add the following:		**Strength of the evaluation's provisions for PROPRIETY:**
Number of Excellent ratings (0-7)	6 x 4 = 24	**[X] 26 (93%) to 28: Excellent**
Number of Very Good (0-7)	1 x 3 = 3	[ ] 19 (68%) to 25: Very Good
Number of Good (0-7)	0 x 2 = 0	[ ] 12 (43%) to 18: Good
Number of Fair (0-7)	0 x 1 = 0	[ ] 5 (18%) to 11: Fair
Total score:	= 27	[ ] 0 (0%) to 4: Poor
		27 ÷ 28 = .964 x 100 = 96%

Accuracy

*THE **ACCURACY STANDARDS** ARE INTENDED TO ENSURE THAT AN EVALUATION EMPLOYS SOUND THEORY, DESIGNS, METHODS, AND REASONING IN ORDER TO MINIMIZE INCONSISTENCIES, DISTORTIONS, AND MISCONCEPTIONS AND PRODUCE AND REPORT TRUTHFUL EVALUATION FINDINGS AND CONCLUSIONS.*

A1 Justified Conclusions and Decisions. [*Evaluation conclusions and decisions should be explicitly justified in the cultures and contexts where they have consequences.*]

a. [+*] Address each contracted evaluation question based on information that is sufficiently broad, deep, reliable, contextually relevant, culturally sensitive, and valid
b. [+*] Derive defensible conclusions that respond to the evaluation's stated purposes, e.g., to identify and assess the program's strengths and weaknesses, main effects and side effects, and worth and merit
c. [+] Limit conclusions to the applicable time periods, contexts, purposes, and activities
d. [+] Identify the persons who determined the evaluation's conclusions, e.g., the evaluator using the obtained information plus inputs from a broad range of stakeholders
e. [+] Identify and report all important assumptions, the interpretive frameworks and values employed to derive the conclusions, and any appropriate caveats
f. [?] Report plausible alternative explanations of the findings and explain why rival explanations were rejected

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. A1, as appropriate:

a. MRE Scientists must be ever vigilant to ensure they are getting the most valid findings from their vendors.
b. While the conclusions are defensible, the warrants behind each MPI to outcome inference should be thoroughly scrutinized at the outset; and revisited when appropriate.

c.
d.
e. This speaks to the importance of the Evaluation and Research Plan, wherein evaluation inferences, interpretive frameworks and values need to be reported.
f. I wonder how much "alternative explanations" are entertained by MRE, their vendors, and the Program Team?

A2 Valid Information. [*Evaluation information should serve the intended purposes and support valid interpretations.*]

a. [+] Through communication with the full range of stakeholders develop a coherent, widely understood set of concepts and terms needed to assess and judge the program within its cultural context
b. [+*] Assure–through such means as systematic protocols, training, and calibration–that data collectors competently obtain the needed data
c. [+*] Document the methodological steps taken to protect validity during data selection, collection, storage, and analysis
d. [+*] Involve clients, sponsors, and other stakeholders sufficiently to ensure that the scope and depth of interpretations are aligned with their needs and widely understood
e. [-] Investigate and report threats to validity, e.g., by examining and reporting on the merits of alternative explanations
f. [+] Assess and report the comprehensiveness, quality, and clarity of the information provided by the procedures as a set in relation to the information needed to address the evaluation's purposes and questions

[ ] 6 Excellent [X] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. A2, as appropriate:

a.
b. I did not inquire into this topic, but I assume it is done by MRE.
c. I did not inquire into this topic, but it is an important one and should be reviewed by MRE team.
d. Program Team reviews and comments on MPERs. I also suspect other stakeholders review and comment on MPERs.
e. I get the sense that this is a tricky subject, but important to support the claims of market transformation.
f. Assume standard MRE processes.

A3 Reliable Information. [*Evaluation procedures should yield sufficiently dependable and consistent information for the intended uses.*]

a. [+*] Determine, justify, and report the needed types of reliability—e/g., test-retest, findings from parallel groups, or ratings by multiple observers—and the acceptable levels of reliability
b. [+*] In the process of examining, strengthening, and reporting reliability, account for situations where assessments are or may be differentially reliable due to varying characteristics of persons and groups in the evaluation's context
c. [+*] Assure that the evaluation team includes or has access to expertise needed to investigate the applicable types of reliability
d. [+] Describe the procedures used to achieve consistency
e. [+*] Provide appropriate reliability estimates for key information summaries, including descriptions of programs, program components, contexts, and outcomes
f. [+] Examine and discuss the consistency of scoring, categorization, and coding and between different sets of information, e.g., assessments by different observers

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comments re. A3, as appropriate:

a. Not sure if this is done.
b. This seems like it should be a part of the MRE Scientist work, but I do not have information on it to make a claim.
c. I think the expertise exists in the MRE department and they should be applied regularly.
d. Not applicable
e. Ditto
f. See MPER5 HPWH

A4 Explicit Program and Context Descriptions. [*Evaluations should document programs and their contexts with appropriate detail and scope for the evaluation purposes.*]

a. [+*] Describe all important aspects of the program—e.g., goals, design, intended and actual recipients, components and subcomponents, staff and resources, procedures, and activities—and how these evolved over time
b. [+*] Describe how people in the program's general area experienced and perceived the program's existence, importance, and quality

c. [+] Identify any model or theory that program staff invoked to structure and carry out the program
d. [+] Define, analyze, and characterize contextual influences that appeared to significantly influence the program and that might be of interest to potential adopters, including the context's technical, social, political, organizational, and economic features
e. [+] Identify any other programs, projects, or factors in the context that may affect the evaluated program's operations and accomplishments
f. [+*] As appropriate, report how the program's context is similar to or different from contexts where the program is expected to or reasonably might be adopted

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. A4, as appropriate:

a. This can be seen in the product and program plans.
b. This seems like a prescription for a type of data on the program and I'm not sure it is appropriate.
c.
d.
e.
f. This is an opportunity to offer program transferability information, but not sure it deserves to be marked absent if not there.

A5 Information Management. [*Evaluations should employ systematic information collection, review, verification, and storage methods.*]

a. [+] Select information sources and procedures that are most likely to meet the evaluation's needs for accuracy and be respected by the evaluation's client group
b. [+] Ensure that the collection of information is systematic, replicable, adequately free of mistakes, and well documented
c. [+*] Establish and implement protocols for quality control of the collection, validation, storage, and retrieval of evaluation information
d. [+*] Document and maintain both the original and processed versions of obtained information
e. [+*] Retain the original and analyzed forms of information as long as authorized users need it
f. [+*] Store the evaluative information in ways that prevent direct and indirect alterations, distortions, destruction, or decay

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. A5, as appropriate:

a. Assumed
b. Assumed
c. Because MRE works with vendors, it may be difficult to implement protocols, but there may be efforts MRE can make towards standardizing expected quality.
d. Again, because MRE works with vendors, the original data are likely not given to MRE until after the processing has been done. Raw data may or may not come to MRE.
e. Ditto
f. Ditto

A6 Sound Designs and Analyses. [*Evaluations should employ technically adequate designs and analyses that are appropriate for the evaluation purposes.*]

a. [+] Create or select a logical framework that provides a sound basis for studying the subject program, answering the evaluation's questions, and judging the program and its components
b. [+] Plan to access pertinent information sources and to collect a sufficient breadth and depth of relevant, high quality quantitative and qualitative information in order to answer the evaluation's questions and judge the program's value
c. [+] Delineate the many specific details required to collect, analyze, and report the needed information
d. [+] Develop specific plans for analyzing obtained information, including clarifying needed assumptions, checking and correcting data and information, aggregating data, and checking for statistical significance of observed changes or differences in program recipients' performance
e. [+*] Buttress the conceptual framework and technical evaluation design with concrete plans for staffing, funding, scheduling, documenting, and <u>metaevaluating the evaluation work</u>
f. [+*] Plan specific procedures to avert and check for threats to reaching defensible conclusions, including analysis of factors of contextual complexity, examination of the sufficiency and validity of obtained information, checking on the plausibility of assumptions underlying the evaluation design, and <u>assessment of the plausibility of alternative interpretations and conclusions</u>

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. A6, as appropriate:

a.

b.
c.
d.
e. There is no metaevaluation, per se.
f. No assessment of alternative interpretations.

A7 Explicit Evaluation Reasoning. [*Evaluation reasoning leading from information and analyses to findings, interpretations, conclusions, and judgments should be clearly and completely documented.*]

a. [+*] Clearly describe all the assumptions, criteria, and evidence that provided the basis for judgments and conclusions
b. [+] In making reasoning explicit, begin with the most important questions, then, as feasible, address all other key questions, e.g., those related to description, improvement, causal attributions, accountability, and costs related to effectiveness or benefits
c. [+*] Document the evaluation's chain of reasoning, including the values invoked so that stakeholders who might embrace different values can assess the evaluation's judgments and conclusions
d. [?] Examine and report how the evaluation's judgments and conclusions are or are not consistent with the possibly varying value orientations and positions of different stakeholders
e. [-] Identify, evaluate, and report the relative defensibility of alternative conclusions that might have been reached based on the obtained evidence
f. [+] Assess and acknowledge limitations of the reasoning that led to the evaluation's judgments and conclusions

[ ] 6 Excellent [ ] 5 Very Good [X] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. A7, as appropriate:

a. With the MPERs, embedded within the life cycle of the product, there is adequate description of the basis for judgments and conclusions. Some of this could be made more explicit by regularly updating an Evaluation and Research Plan.
b.
c. "values invoked" is a difficult one to respond to
d. I don't see this being documented, but it is being done internal to MRE. So, no reporting.
e. No consideration of alternative conclusions.

f. This is ideally done early in the development of MPERs for measuring MPIs, but surely it must be done at the judgment moments, such as the move into long term monitoring and tracking.

A8 Communicating and Reporting. [*Evaluation communications should have adequate scope and guard against misconceptions, biases, distortions, and errors.*]

a. [+] Reach a formal agreement that the evaluator will retain editorial authority over reports
b. [+] Reach a formal agreement defining right-to-know audiences and guaranteeing appropriate levels of openness and transparency in releasing and disseminating evaluation findings
c. [+] Schedule formal and informal reporting in consideration of user needs, including follow-up assistance for applying findings
d. [+*] Employ multiple reporting mechanisms, e.g., slides, dramatizations, photographs, powerpoint©, focus groups, printed reports, oral presentations, telephone conversations, and memos
e. [+] Provide safeguards, such as stakeholder reviews of draft reports and translations into language of users, to assure that formal evaluation reports are correct, relevant, and understood by representatives of all segments of the evaluation's audience
f. [+] Consistently check and correct draft reports to assure they are impartial, objective, free from bias, responsive to contracted evaluation questions, accurate, free of ambiguity, understood by key stakeholders, and edited for clarity

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. A8, as appropriate:

a.
b.
c.
d. This seems non-applicable, but likely MRE Scientists and NEEA leaders provide various methods of communicating MPERs and other evaluation findings.
e.
f.

Scoring the Evaluation for ACCURACY Add the following:	**Strength of the evaluation's provisions for ACCURACY:**
Number of Excellent ratings (0-8) 5 x 4 = 20	[ ] 29 (91%) to 32: Excellent
Number of Very Good (0-8) 2 x 3 = 6	**[X] 21 (66%) to 28: Very Good**

Number of Good (0-8) 1 x 2 = 2 Number of Fair (0-8) 0 x 1 = 0 Total score: = 28	[ ] 13 (41%) to 20: Good [ ] 5 (16%) to 12: Fair [ ] 0 (0%) to 4: Poor **28 ÷ 32 = .875 x 100 = 88%**

Evaluation Accountability

*THE **EVALUATION ACCOUNTABILITY CRITERIA** ARE INTENDED TO ENSURE THAT AN EVALUATION IS SYSTEMATICALLY, THOROUGHLY, AND TRANSPARENTLY DOCUMENTED AND THEN ASSESSED, BOTH INTERNALLY AND EXTERNALLY FOR ITS UTILITY, FEASIBILITY, PROPRIETY, AND ACCURACY.*

E1 Evaluation Documentation. [*Evaluations should fully document their negotiated purposes and implemented designs, procedures, data, and outcomes.*]

Document and preserve for inspection the following:

a. [+] Contract or memorandum of agreement that governed the evaluation
b. [+] Evaluation plan, including evaluation tools and resumes of key evaluation staff
c. [+] Evaluation budget and cost records
d. [+] Reports, including interim and final reports, the evaluation's internal metaevaluation report, and, if obtained, a copy of the external metaevaluation report
e. [+*] Other information determined to be needed by reviewers, such as technical data on the employed evaluation tools, a glossary of pertinent theoretical and operational definitions involved in the evaluation, a description of the subject program, a record of stakeholder involvement, and news accounts related to the evaluation
f. [+*] Evidence of the evaluation's consequences, including stakeholders' uses of findings

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. EA1, as appropriate:

a. This section likely has two foci: the evaluation of the MT and the individual MPERs. The MPIs and the assorted other documentation, like logic model, program and product plans, and MT story all document the information needed to undertake evaluating the MT. The one piece that is not fully incorporated is the Evaluation and Research Plan.
b. While MPER plans are well-documented, the overall Evaluation and Research Plan is less formalized across programs.
c.
d. There is no metaevaluation effort to be reviewed, but all reports are made public and shared widely.

e. This would be embedded in the Evaluation and Research Plan. This is a clear opportunity for MRE to improve. Take ownership as the knowledge broker that is keep records of program logic, outcomes, MPIs, and evaluation against them.
f. Ditto

E2 Internal Metaevaluation. [*Evaluations should use these and other applicable standards to examine the accountability of the evaluation design, procedures employed, information collected, and outcomes.*]

a. [-] At the evaluation's beginning, determine the metaevaluation's intended users and uses (e.g., formative and summative)
b. [-] Develop a plan for obtaining, processing, and reporting a sufficient scope and depth of information to assess the evaluation's utility, feasibility, propriety, and accuracy and address the intended users' needs for timely metaevaluation feedback and reports
c. [+] Assign responsibility for documenting and assessing the evaluation's plans, process, findings, and impacts and budget sufficient resources to carry out the internal metaevaluation
d. [-] Maintain and make available for inspection a record of all internal metaevaluation steps, information, analyses, costs, and observed uses of the metaevaluation findings
e. [-] Reach, justify, and report judgments of the evaluation's adherence to all of the metaevaluation standards
f. [-] Make the internal metaevaluation findings available to all authorized users

[ ] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [X] 0-1 Poor

Comment re. EA2, as appropriate:

a. This may be embedded in the MRE process for the MPER contracts, but it was not clear to me.
b. This may be embedded in the work done with the Program Team, but it was not clear to me.
c. The regular review of RFPs is a form of internal formative metaevaluation.
d. I saw no evidence of this
e. Ditto
f. Ditto

E3 External Metaevaluation. [*Program evaluation sponsors, clients, evaluators, and other stakeholders should encourage the conduct of external metaevaluations using these and other applicable standards.*]

a. [+*] Confirm through exchange with key stakeholders the need for an external assessment of the evaluation and the purposes it should serve (e.g., formative or summative)
b. [+] Stipulate that these and possibly additional standards will be used to assess and judge the evaluation
c. [+] Select, recruit, and reach a formal agreement with an external metaevaluator who possesses an independent perspective, appropriate expertise, and freedom from possibly compromising connections or interests
d. [+] Assure that the external metaevaluation is adequately planned, staffed, and funded
e. [+] Provide the external metaevaluator with access to information and personnel required to conduct a thorough, defensible metaevaluation that serves the intended purposes
f. [+] Assure that the metaevaluation will be subjected to appropriate quality control and that the metaevaluator will deliver as part of the metaevaluation report an attestation of its adherence to the metaevaluation standards

[X] 6 Excellent [ ] 5 Very Good [ ] 4 Good [ ] 2-3 Fair [ ] 0-1 Poor

Comment re. EA3, as appropriate:

a. Assuming the current review is the external metaevaluation for MRE's MT evaluations.

Scoring the Evaluation for EVALUATOR ACCOUNTABILITY		**Strength of the evaluation's provisions for EVALUATOR ACCOUNTABILITY:**	
Add the following:		[ ] 11 (92%) to 12:	Excellent
Number of Excellent ratings (0-3)	2 x 4 = 8	**[X] 8 (67%) to 10:**	**Very Good**
Number of Very Good (0-3)	0 x 3 = 0	[ ] 5 (42%) to 7:	Good
Number of Good (0-3)	0 x 2 = 0	[ ] 2 (17%) to 4:	Fair
Number of Fair (0-3)	0 x 1 = 0	[ ] 0 (0%) to 1:	Poor
Total score:	**= 8**	**8 ÷ 12 = .667 x 100 = 67%**	

D. Components of Argumentation Logic

Warrants

As the figure in the text above shows, warrants legitimate the inference from evidence to claim. What this means is that warrants give authority to why the evidence is in fact what is needed to make a claim. For instance, if one wanted to claim that some training given to installers made them more likely to recommend HPWH, one might point to evidence of before and after training data for what type of water heaters the installer installed. The warrant for this claim could be other research that shows installers more often install water heaters they are most comfortable with installing. The warrants "reveal that the reasoning step between evidence and claim is legitimate" (Fournier & Smith, 1993) and is the first line of support for the evidence-to-claims link.

Backings

Backings call in greater support for warrants, when needed. Some warrants are weaker than others, leaving room for doubt. For instance, in the example just above, the backing for this warrant could be that installers are surveyed before training to learn of their preferences and finds that pre-training most installers are not aware of HPWH, therefore, change in knowledge correlates with installer data.

Qualifiers

When evidence is martialed in support of a particular claim, it often has limitations. These can be referred to as qualifiers because they, quite literally, qualify the likely strength of any such claim. In the above example, the before- and after-training data that 'shows' that installers install more HPWH that they are trained to install is only as strong as the response rate (coverage) of the installers. The claim that the training influences installers to install more HPWH is likely only as strong as the reporting by installers.

Conditions

Conditions can be thought of as exemptions. When some evidence is presented to support some claim, there are times when this evidence might not be entirely trusted. For instance, if we to find out that installers were also receiving incentives to install HPWH then our claim that the training influenced their choices is highly suspect.

The value of having these concepts laid out can be seen in the day-to-day work MRE Scientists do with the Program Team. Interrogating claims is the job of MRE Scientists. For example, when selecting an MPI the MRE Scientist critiques the premise that it is in fact good evidence of an outcome. One claim that should likely be interrogated on a regular basis is the idea that the MPIs satisfactorily constitute a "preponderance of evidence" as a proxy for indicating market progress towards transformation.

E. NEEA MRE Assessment Interview Protocol

Semi-structured protocol guiding the interview

This protocol will be used to ensure that all relevant discussion points are addressed during the interviews with NEEA personnel. It is semi-structured, in that the questions are set out in the below order but there is no specific need to ask them in this order. If the conversation is fluid and some later topics arise before others, then they will be pursued as naturally occurring. The interviewer should return to this list of questions often to ensure to not miss a topic. Ideally, in the end, all questions will be addressed.

This protocol is organized to be useful for all interviewees. Where particular items are intended only for a subset of interviewees, these are called out with a name in [brackets] at the end of the item. First the welcome script and then the questions.

Welcome Script

Thank you for your time, I'm doing an assessment of MRE's work, focusing on two concepts. First, I hope to describe their work in relation to existing evaluation knowledge. Secondly, I hope to use that analysis to raise awareness of the strengths and any opportunities for improvement or growth by the department and perhaps for the organization more generally.

In particular, in our discussion, I would like to learn more about MRE by discussing the evaluation of NEEA's Market Transformation programs.

Guidelines

Just a few points to note before we get started.

I would like to record this interview, for notetaking purposes. I can generally take verbatim notes when I am not doing the interview, otherwise, my notes will be rather incomplete. Recording will take the guesswork out of my memory and limit my needing to come back and ask you to verify something I thought I heard you say.

Are you okay with me recording this interview for notetaking purposes only? **[if yes, start the recording]**

To that point, while it would be difficult to offer full confidentiality, I will endeavor to characterize my learning in ways that does not point to specific

voices. If there is something that you would like to share with me that you would prefer not to get back to you, please point that out so I can make extra effort at confidentiality.

Along those lines, the only people at NEEA that are aware of who I'm interviewing are Amy and Susan. The exception to that is the MRE Scientists, who will know that I am interviewing some of their department colleagues, just not which ones.

Finally, if I do not cover something that you think is important, please let me know.

Do you have any questions for me before we get started?"

Questions

The questions below are organized in a logical order from broad to more specific, concluding with a section for only the MRE Scientists and their Director.

General

1. Please tell me your job title and your role in NEEA's market transformation programs? [All]
2. How would you describe the role of the MRE department, specifically for NEEA's MT programs? [All]
    a. *Though I have read quite a bit in reports, RFPs, and presentation materials, I would like to hear it in your own words; Perhaps use a recent or current project to describe MRE's role from your perspective*.

Interactions Between Departments

3. How do you, in your position at NEEA (and others in your department, e.g., market analysts more generally), interact with MRE Scientists? [All except MRE Scientists]
    a. *Perhaps use a recent project as an example to describe the interaction*
    b. Probe: How does MRE support your role at NEEA? [All except MRE Scientists]

Strengths/Weaknesses

4. What have you experienced as the biggest strength(s) of MRE? [All]
5. Any weaknesses that you have noticed? [All]

Specific Questions

1. How well do you think the current model of theory-informed evaluation (i. e., market characterization, theory of change, logic models, MPERs, and MPIs) serves NEEA? [All]
    a. Probe: How could it be improved?
2. Assuming there are critics of MRE, what do (or would) they say? [MRE Scientists]
    a. Probe: How could MRE Scientists improve their contribution to evaluating NEEA's market transformation programs?
3. Where could MRE Scientists' capacity be improved? [All]
    a. Probe: How much interest is there in building the capacity of all NEEA MT-focused staff to become stronger methodologists? [All]
    b. Probe: What capabilities do you think all MRE Scientists should have? [MRE Scientists…others]
4. What did I not ask you that you think I should have? [All]