Using Program Theories to Align Performance Metrics with Public Purpose Goals

Jeff Erickson, PA Government Services Inc.
Oscar Bloch, Wisconsin Department of Administration

ABSTRACT

Wisconsin is implementing statewide energy efficiency programs with public benefits funds. The programs are being run through the state government and are being implemented and evaluated by private firms. The programs must address a variety of legislative goals including energy efficiency, system reliability, environmental protection, and rural economic development. At the same time, the program implementers must establish a portfolio of programs that balance market transformation and resource acquisition policy objectives. With such a range of issues to address, the process of designing effective and responsive programs and evaluations has taken substantial effort. Under direction from the state, the contractors were instructed to take a theory-based approach to program and evaluation design. However, the programs were quickly designed and fielded based on incompletely articulated program theories. As a result, substantial effort has been expended attempting to develop viable program theories for second-year contracts. The state has announced that key metrics in the program theories will be incorporated into second-year contracts to ensure that the implementers' motivations are in line with the broader goals. This paper will describe the methods used to facilitate discussions between the state, evaluators, and program staff to develop and refine program theories that are consistent with performance metrics and overall public policy. It will describe the key ingredients needed for effectively supporting the development of viable and useful program theories. It will also present some guiding principals for judging whether a program theory is adequate or not. Along the way, the paper will present the pitfalls encountered, approaches to solving those pitfalls, and lessons learned from the effort.

Introduction

In the summer of 2001, Wisconsin began implementing statewide energy efficiency programs with public benefits funds, replacing programs previously run by investor owned utilities under Public Service Commission regulatory supervision. The programs are being run under the banner of Wisconsin's Focus on Energy program. The programs are being run through the state Division of Energy in the Department of Administration (DOA) and are being implemented and evaluated by private firms. Programs are combined in three main groups, Residential, Major Markets (commercial and industrial), and Renewables. Each group is the responsibility of a single, private sector administrator and their team of subcontractors. Each group has multiple programs, typically addressing distinct market sectors. The Major Markets effort includes programs addressing industrial, water and wastewater, small retail, schools, new and existing buildings, and agriculture. Residential programs include new construction, Energy Star products, low income, and multi-family. The programs must address a variety of legislative goals including energy efficiency, system

reliability, environmental protection, and rural economic development. At the same time, administrators must establish a portfolio of programs that balance market transformation and resource acquisition policy objectives. With such a range of issues to address, the process of designing effective and responsive programs and evaluations has taken substantial effort. This paper will describe the road taken in Wisconsin, the successes and failures we encountered, and the lessons we learned from the process.

DOA recognized the importance of theory-based programs and evaluation and explicitly required theory-based approach in its request for proposals for the Major Markets contract. (Since the implementation methods vary significantly between some programs this implies separate theories for each program or at least for groups of programs that employ similar methods.) The Request for Proposal (RFP) stated, in part, "The Major Markets Administrator must provide, first of all, a clear elucidation of the theory underlying its program design, including target markets and barriers, anticipated outcomes based on the theory and how its programs are consistent with that theory."

The Major Markets Administrator's first year contract also contains language stating that the programs will be designed using a theory-based approach. The RFP and contract also specify resource acquisition (energy savings) and market effects goals, as does DOA's policy statement on balanced portfolios. The program's first year contractual requirements for resource acquisition specify a measurable target (achieving a cost/benefit ratio of one) but not so for the market effects requirements. They specify defining a means to identify, document, and value the benefits from market effects, but do not include specific targets.

Although the RFP and contract called for a central role for program theories, the reality fell some distance short of that goal. The program plans attached to the Major Markets first year contract contained short sections titled "Program Theory" but most were essentially a list of barriers and program actions intended to target those barriers. End goals were not spelled out clearly and the logic for how program actions would achieve results was not evident. Mid-way through the first contract year, program and evaluation staff worked together to attempt to clarify the program theories. Some program theories were significantly improved through that process but the majority continued to have significant problems with logic, comprehensiveness, and structure. Part way through year 1, DOA announced their intention of including market effects metrics in year 2 contracts to specifically address market effects goals. This lead to a renewed attempt on the part of program and evaluation staff to improve the program theories. As this paper is being written, DOA and the administrators are negotiating second year contracts that will include market effects metrics to specifically address market effects goals.

In the process of trying to achieve agreement on theories and metrics, DOA, program staff, and evaluators have encountered many successes and failures. Understanding the many factors that can stand in the way of creating viable program theories and contractual theory-based metrics can help program designers, funders, and evaluators plan ahead and avoid some of the pitfalls encountered in Wisconsin. It is not enough to specify theory-based programs in RFPs and contracts. To ensure success, all parties (funders, administrators, and evaluators) need to work together to bring the theories and processes into being. Developing a well-designed program theory that supports contractual metrics and broader evaluation goals can be quite difficult. Neither the process of developing the theory nor the theory itself necessarily needs to be complex. However, it is quite important that no one underestimate the effort and time needed to achieve a well-designed theory.

Criteria for Program Metrics

Even with a good theory, determining which metrics should be incorporated in a contract can be difficult. Such metrics should be timely, theory-critical, and measurable. Many metrics that fit one of these criteria may not fit the others. There are many factors that have to be balanced, including the following.

Timely

Is it likely that the effect will be large enough to be measurable in the contract year? How likely is it that the margin of error will include zero? If a metric is going to be incorporated in a program contract and if, as in Wisconsin, the contract period is one year, then the program theory must show plausible connections between incremental changes in metric outcomes annually and the ultimate outcome predicted by the theory. This points to the critical role of the program theory: since the ultimate indicators of market transformation can take years to appear, contractual metrics must be pre-cursors to the ultimate effect and the program theory is the logical place to spell out what those pre-cursors are and how they will lead to the ultimate effects.

Theory-Critical

How central to the theory is the metric? How strongly is the metric correlated with ultimate effects? Does the metric form one of the key links between program action and ultimate effect? There will always be a temptation to go after the "low-hanging fruit" – the metrics that are clear and easily measurable. However, if those metrics do not have a strong role in the program theory, their value for predicting the ultimate effects of the program will be weak and their value as contractual metrics will be low. If, for example, lack of time to devote to learning about and implementing energy efficient projects was the critical barrier in a particular market and the program was designed to provide services to reduce time commitments and to provide financing, a metric addressing the financing aspect of the program would be much less useful than one addressing the time-related services.

Measurable

The chosen metrics should meet several criteria to ensure that they can be reliably and consistently measured, including the following:

- The chosen metrics should be clear and unambiguous to minimize the possibility of misinterpretation when designing measurement methods and analyzing results.
- The chosen metrics should present a reasonable possibility that they can be measured with precision at an acceptable cost. This might rule out metrics that can be measured precisely but only at high cost (this could include, for example, sales data for some types of equipment), as well as those whose very nature makes us suspect that they can never be measured very accurately (perhaps customer attitudes or knowledge).

- The chosen metrics should be amenable to data collection and reporting methods that can be open, visible, and easily verified or double-checked to reduce the likelihood that the program administrator could manipulate the data undetected.

Net or Gross

Should the contract be based on unadjusted data collected and reported by program staff? Or should it be based on evaluation-adjusted results? (Should it be based on gross or net values in impact evaluation terms?) Contracts based on net results will provide motivation to the program to minimize the factors that reduce gross (e.g., free ridership). However, contracts based on gross results will probably be easier to negotiate and be more predictable.

Facilitating Theory Definition and Metric Selection

If creating complete and useful program theories and identifying market effects metrics for near-term contract monitoring is difficult, then what can be done to help the process? In Wisconsin, DOA staff, evaluators, and program staff undertook a number of activities to develop contractual market effects metrics. Those activities can be placed in four groups:

- Motivate
- Inform
- Empower
- Assist.

Motivate

Without motivation, program staff will not care enough to create and use a complete, well-designed program theory. Similarly, evaluation staff will not care enough to design an evaluation around the program theory. Efforts to motivate staff are in a sense either internal or external, representing the carrot and the stick.

Internal—the carrot. Carrot-based approaches to motivation help implementers and evaluators see the value and self-interest in creating and using complete program theories. Many new efforts to create program theories would get off on the right foot if they begin with a training session on program theory design. The first topic in that session should demonstrate how program theories can help improve program design, delivery, evaluation, and performance. It should show the trainee how creating a program theory will make their life easier, not more difficult, and their work more effective and productive. Only if people are properly motivated will they apply themselves to learning and applying knowledge about program theories.

Our experience in Wisconsin demonstrates how important it can be for program implementers to believe in the usefulness of program theories to meeting their goals. Work on program theory in the Focus on Energy (Focus) program began with joint working sessions between evaluators and implementers in the pilot program, before the statewide

program began. They took an experimental, learn-as-we-go approach developing theory display templates to be compatible with existing program task descriptions and accepting a fairly simple, short explanation of the theory. The results were sufficient for the initial approaches in the pilot but the program staff approached their theories more as a component of their reporting requirements than as an integral and useful part of program design and management. Work on program theory in the statewide program also began as working sessions between evaluators and program staff. The approach chosen for describing theories and the rigor demanded of the logic in the theories was significantly improved from the pilot but most program staff continued to treat their theories as a reporting issue, not a program design and management tool. In an attempt to break an apparent deadlock, DOA brought in an independent, outside program theory expert for a day of training. The training began with descriptions and examples of the usefulness of program theories. Subsequent work on program theories indicates that some program managers are now program theory converts. However, others continue to approach the effort as required reporting. This experience has demonstrated the importance of obtaining meaningful buy-in from all involved on the importance of program theories and their potential benefit to the successful design and management of programs. If the internal motivation approach fails, however, it may still be necessary to provide external motivation, as the next section discusses.

External—the stick. Stick-based approaches will probably utilize contractual requirements and penalties. The funding organization can insist that a theory-based approach be used in the program design phase. To be credible, the requirements and consequences of noncompliance should be clear up front. In Wisconsin, the Major Markets RFP and contract specified a theory-based approach. However, as we discussed above, the letter of the law was met (the program implementers created program theories) but the spirit of the law was not since the theories were incompletely defined and in some cases logically deficient.

Two approaches could help avoid this outcome. First, the RFP could request that proposals outline a program theory (it is probably too much to expect a fully developed theory in a proposal) and propose an approach for defining programs using program theories. The proposal scoring then should give significant weight to the quality of the program theory approach. Second, the RFP should specify a timeline that incorporates program theory development as an explicit stage at the beginning of program design work. Rushing to the field with programs and then trying to develop theories after the fact will only make it harder to create and use program theories.

However, if the program is already designed and running, the funding organization may still be able to exert pressure to motivate program staff to create and use program theories. In Wisconsin, DOA has announced its intention to 1) incorporate market effects metrics in second year contracts, and 2) require agreement between DOA program managers, implementers, and evaluators on the essential theory and measurement of its predicted outcomes for each program before final approval of the contract. That process is underway as these words are being written. The pending contract negotiation process, perhaps combined with the internal motivation discussed above, has led some program managers to create theories that are a significant improvement on earlier versions. They report in interviews that they now see the benefit of creating these theories. Time will tell whether they use them as management tools.

Inform

Theory-based program design is not a new concept in the business world and has been tried for a few years in some parts of the country, however it is relatively new to energy efficiency program design in many places, particularly in Wisconsin and has not been effectively and thoroughly implemented in many places. As a result, it is probably a good assumption that any program theory design effort should begin with a training session. Our experience in Wisconsin indicates that three of the most important things the training should focus on are the following:

- The difference between program barriers and market barriers, between program outputs and outcomes, and their position in the program theory.[1]
- The necessity that the program theory specifies a sequence of specific, causally linked market effects that lead all the way from program actions to ultimate effects.
- Providing the appropriate level of detail – enough to make the sequence of effects clear but not so much that the structure gets lost in the detail and that developing the theory takes too much time and effort.

In Wisconsin, training has taken several forms. It began as formal and informal presentations by evaluators to program staff. It was extended through one-on-one sessions as evaluators and program staff worked through the details of specific program theories. Some of these sessions proved productive as program theories were clarified and evaluators and program staff alike saw their usefulness. However, other sessions proved frustrating for evaluators and program staff alike as they went through multiple rounds of comment and review with little evidence of a meeting of the minds. Because of the lack of significant progress in some areas, we decided that it was time to bring in outside help to provide a more formal training effort. It was hoped that an independent expert would be more convincing since they would not have vested interests in the outcome and that they would be able to provide practical advice for moving theories forward. Subsequently, DOA brought in an independent program theory expert to give a one-day course on program theory for DOA, program, and evaluation staff. The course provided a framework for laying out a program theory and recommendations for steps to take to ensure that theories are complete and logically consistent. The course was well received and met the second goal mentioned above, providing practical advice. The independence of the expert was somewhat less successful than we hoped in convincing skeptics that program theories were worthwhile. As a result, some successfully created viable theories but others continued to struggle.

To further help theory authors, DOA developed a set of goals and metrics guidelines to assist them in the development of their program theories and the specific goals and metrics in those theories. The guidelines provided objectives, definitions of terms, directions for types of metrics to develop, and an overview of the contract negotiation process to help people zero in on the key goals and metrics that could be included in their contracts.

---

[1] Program barriers and program outputs refer to issues of implementation, market barriers, and market outcomes relate to the hoped-for ultimate results in the market. In a sense, program barriers are to program outputs what market barriers are to program outcomes. According to one source, outputs are the activities, products, methods, services that reach people and users. Outputs lead to Outcomes: the results and benefits for individuals, groups, agencies, communities and/or systems" (UWEX 2002).

Empower

Motivation and education alone are not enough to ensure programs are designed on complete program theories. Program managers and designers must still be empowered to create those theories. Three steps can help:

Allocate time. Program theory design cannot be treated as an afterthought or a sidelight and be successful. Time should be specifically allocated, by contract if possible, for program theory development in the program design phase. If the program theory is being developed after a program has already hit the field, it is still necessary to ensure that enough time is dedicated to the effort and with the right emphasis. If the staff feel that the design time is stolen from their "real" job of delivering services, the theory will suffer. One resolution to this dilemma is to look carefully at who is in the best position to design the theory – it may be that someone more senior with a good grasp of the entire effort but without day-to-day management responsibilities would be in a better position to devote time to the effort. This is closely related to the following step.

Allocate staff. Program theory design is not rocket science but neither is it akin to a jigsaw puzzle, where senior staff could give pieces to junior staff and expect that they can make the picture clear. It is important that staff with substantial knowledge of the program goals and experience with program design tackle the program theory. The danger in using staff inexperienced in program design to develop required program theory is that the product may meet the "letter of the law" without coming close to "the spirit of the law." It is important that people with the power to make decisions are intimately familiar with the concept of program theory and have committed resources to development and testing of the theory if it is to be used as an effective tool for program design. This was a particularly significant problem in Wisconsin. Some of the staff initially assigned to developing the program theory were either too junior to have a good grasp of the issues or too intimately involved in keeping the programs running to be able to devote enough time and interest to the task. When DOA recognized this problem they encouraged the contractor's management to re-arrange staff assignments to ensure that the theories got the attention they needed from the right staff. Often this meant assigning senior management staff to the effort – those responsible for general program direction but not for day-to-day management.

Coordinate structure. There are many possible ways of approaching the design and presentation of program theories. Some approaches may work better for specific kinds of energy efficiency programs than others. If the effort at hand is a single program, then the program designers can choose the approach that seems to fit their situation best. If, on the other hand, multiple programs are being implemented under one umbrella, as is the case in Wisconsin, then choosing an approach that can be used by all programs may have benefits.

While there may be separate teams at the field level, there will often be fewer divisions at higher management levels, including among the funding agencies and evaluators. The Wisconsin programs exhibit a definite pyramid-shaped hierarchy – while there are many unique programs, there are only a handful of people responsible for the program theories. For this reason, a single approach was used for presenting each program theory. The earliest theory efforts used a basic framework or template that was designed by the evaluation team,

at the request of the program team. During the formal program theory training, a simpler framework was presented and it was adopted for subsequent work.

Assist

Motivation, education, and empowerment set the stage for effective program theory design. To complete the play, you need to get all the actors on the stage and then help them work together. Two aspects need attention: the process and the details.

The process. Who is going to work on the program theory and how are the different team members going to work together to achieve their goals? In most cases, the program theory needs to be the primary responsibility of the program implementation team. Only then will they embrace it with enough enthusiasm to ensure that it is well developed and used for program design and fine-tuning. That said, however, it is also probably important to involve other organizations in the process for two reasons. First, since the program theory must be understandable to and meet the needs of the program, funding organization, and evaluators, they should be involved. Second, a theory designed in relative isolation can end up being narrowly focused and, perhaps, not as thorough as it should be. There can be many, many possible program impact pathways in even relatively simple programs. Different perspectives can help identify the most important or influential pathways and thus improve the theory.

In Wisconsin, the program managers worked with evaluation team members to define and fine-tune the theories. Evaluation staff provided suggestions for presentation and critiques of theory completeness and layout while program managers were responsible for the program theory content and logic. As we discussed above, this process worked well in some cases but seemed to enter an endless cycle of review, instruct, and rewrite in other cases when evaluators and program staff could not see eye-to-eye. DOA called in the independent expert to attempt to break this cycle. Again, that worked in some cases but not in others.

The details. Finally, if everything described above has been attended to, you hopefully will have a motivated, knowledgeable team working together to create a program theory. How can you tell whether their product is a complete program theory or a weak imitation? How can you ensure that their explanation of that theory does it justice? We offer some suggestions below. A thoroughly defined program theory should contain the components listed in the following table.

One of the critical components of a well-defined program theory is an explanation of the causal linkages between interrelated market effects. It may be that one of the program's ultimate goals is increased saturation of energy efficiency goods and services but its primary methods of achieving that goal do not directly create the changes that define that goal. Rather they initiate one of the first steps toward that goal, say increased awareness of energy efficiency services or a change in attitudes about energy efficiency products. This first-order market effect (changed attitudes) in turn creates the ultimate market effect of increased purchases. The program theory should clearly define the first-order market effects and subsequent market effects created by the first order market effects. It is very important that the theory shows the chain of events that the theory postulates starting from the initial, direct impact and leading to the ultimate effect. It is through these causal sequences that the theory will show how the program will achieve broader and longer-lasting market effects.

Table 1. Program Theory Components

Column Heading	Definition
Specific Program Activity	The specific actions that the program will implement that are intended to affect the market actors and create the market effects. These should be as specific as possible.
Market Actor Targeted by Activity	The target for the specific program activity and the population where we should see the initial market effects. This should be the decision-maker the program activity is trying to affect. The definition should be as precise as possible whether the target is program allies or participants. For example, if you can associate a particular specific program activity with architects, then the theory should treat that activity and market actor separately rather than lumping the activity with others addressing a general market actor category of "Program Allies."
Market Barrier or Opportunity	The barrier that is being targeted by the specific program activity. The barrier should be specific enough to address what needs to be changed to see a change in behavior in the targeted market actor.
Market Effects	The changes that you expect to see in the specific, defined market actor as a result of implementing a specific activity. Each distinct market effect should be delineated in the theory. First-order market effect (that is the one directly created by a specific program activity) should be clearly identified. Follow-on market effects, which are secondary effects created by the primary market effect, should be identified when they are necessary for showing a sequence of market effects that lead to an ultimate effect.
Timeframe	When the market effects can be expected to occur.

All market effects should be connected to a sequence of market effects that ultimately leads to a key public policy goal of the program, such as energy efficiency, environmental protection, or energy supply reliability. It is important that the ultimate goal of each activity be clearly defined in the theory.

Components of a Complete Program Theory

The components of a complete program theory can be summarized in a relatively few rules, as follow:

- Addresses all relevant market actors influenced or involved (and gives detail so it is clear who the market actors are)
- Describes program stimulus
- Describes market effects expected and the timeframe they are expected in
- Links program stimulus to expected market effects
- Links one market effect to another when one effect causes another (e.g., increased demand causes increased supply)
- Indicates the expected sequence of market effects showing which effects come before which (this is part of the cause and effect linkage)
- Describes the market barriers overcome
- Links the market effects with the market barriers
- Describes the decision maker targeted by the program action and the decision that the program action is intended to affect.

Summary and Conclusion

If program theory is going to be taken seriously and if the goal is to have it play a crucial part in program design and management, then program theory design needs to be undertaken early, with vigor, and with strong backing from those holding the purse strings. Theory design efforts are unlikely to produce viable, useful results if they are treated as a side issue, to be dispensed with as quickly as possible to satisfy evaluator or funding organization demands.

It is critical to attend to each and every issue described in this paper if the effort is to succeed. But even so, this is no guarantee of success. In Wisconsin, we went about attending to these issues in an incremental fashion, tackling first one, then another. While we ultimately did attend to each we still failed (as of this writing) to create viable theories for all programs. Some programs developed strong, logical, and complete theories, however other writers just never seemed to come to grips with the purpose and spirit of the effort, despite extensive training, support, and assistance.

We saw two critical failings. First, some responsible for program theories could not get beyond treating program theory as a summary description of program implementation activities. Their theories spoke of program barriers not market barriers and their "market effects" were program achievements (so many participants, so many training sessions) not true market effects. It was as if they were treating program theory as an exercise in describing the activities they had already decided to implement and the barriers in the way of successfully implementing them. Second, some did see the importance of addressing market barriers and true market effects indicators but their theories were too disconnected from reality to be believable or useful. They wrote their theory to include likely barriers and a grand vision of appealing market transformation results but failed to describe in any convincing manner how their activities could possibly produce the appealing results.

For a variety of reasons, work on program theories consumed substantial resources over a many months. The total time spent on program theories was probably more than if a concerted and focused effort was made early in the program planning process. Some involved in the process of working on the Wisconsin program theories have come to the conclusion that the capabilities typically found in effective program managers (that is day-to-day managers, not necessarily the program designers) may not be those that are needed for creating viable program theories. If the time and resources are not devoted to defining program theories during the program design process then it may be more effective to assign the role of program theory drafting to evaluators or independent program theory experts and then obtain general buy-in from program managers. However, while this route might be a more efficient and cost-effective approach to creating quality program theories it raises the danger that program staff will not feel a sense of ownership or connection with the theory. This could reduce their inclination to use the theory as a management tool and to feel commitment to the goals and metrics embodied in the theory.

Others involved in the process have now come to the conclusion that a more effective approach to program theory design would be to organize extensive working sessions with evaluators and/or program theory experts working with program design staff. The precise details of the program theory should be developed in the working sessions. In Wisconsin, we typically took a draft-comment-redraft approach and failed to achieve consensus in several cases. As these words are being written, DOA staff are planning sessions that will bring

DOA, program, and evaluation staff together to iron out the final details of the program theories and market effects indicators for the contract. Time will tell if that effort proves successful.

References

Jordan, Gretchen. 2002. "Using Logic Models to Define a Balanced Set of Metrics for Performance Contracts." Madison: Wisconsin's Focus on Energy Training.

University of Wisconsin-Extension (UWEX). 2002. http://www.uwex.edu/ces/pdande/Evaluation/logicmodels.htm.