
6

THE MONITORING- EVALUATION- RESEARCH CONTINUUM: INSIGHTS FROM THE MASSACHUSETTS MANUFACTUR- ING PARTNERSHIP

Eric S. Heller

Eric S. Heller is Director of Research and Evaluation with the Maurice A. Donahue Institute at the University of Massachusetts, Amherst, MA.

The establishment of and adherence to unambiguous, commonly agreed upon definitions of central concepts and terms is vital to the conduct of constructive professional discourse. Conversely, the degree to which different discussion participants apply multiple meanings and expectations to critical concepts frequently impedes progress towards group consensus on key issues. In some cases, miscommunication of this type can lead to significant derailing of group processes and skepticism among participants as to one another's credibility.

"Evaluation" is a prime example of a critical term for which a broad range of definitions, interpretations and activities apply. Individuals' subjective experiences may bias their orientation to and perceptions of the evaluation function. Resultant conflicts among individual perceptions often hinder progress in developing a specific evaluation approach and protocol for a program. The dilemma is that while participants in the discussion believe they are discussing a single concept, "evaluation", they may in fact be referring to a broad array of "information needs", covering a continuum that ranges from fundamental project monitoring to controlled experimental research. As the number and types of program stakeholders increase, the potential for conflicting frames of reference in evaluation planning similarly grows.

The types of individual experiences that may lead to this ambiguity are countless. Within the world of federally-funded programs, the term "evaluation" can assume any number of identities and refer to a range of activities differing greatly in objectives and scope. For example, some multi-million dollar grants to research centers require minimally funded "evaluation" components that encompass only perfunctory monitoring and tracking. On the other hand, many smaller, service-oriented demonstration grants stipulate that 15-20% or more of direct project funds be allocated to program evaluation, with requirements that the evaluator provide rigorous process-outcome-impact studies as well as customer satisfaction assessment and ongoing management feedback and support. Managers of programs with such divergent evaluation requirements will almost inevitably have

similarly divergent perceptions of the nature and function of "evaluation."

An individual's personal experience as the subject of "evaluation" may also contribute to communication problems. A professional whose past job may have required periodic performance evaluation, or whose previous project underwent third party evaluation, limited to assessment of the attainment of stated goals, may enter into the evaluation dialogue with negative biases. This negativism may be based on a perception of evaluation as an intimidating, intrusive or judgmental process, focused on identifying personal or programmatic flaws and failings. The evaluator who introduces a comprehensive formative and summative evaluation strategy, focused on internal management support, must overcome this experience-based perception in order to facilitate constructive dialogue, access meaningful data, and achieve true integration of evaluation and program. Clarity of the intention and definition of "evaluation" is critical to productive communication and education.

Within the realm of industrial extension services, the challenge to establish clear-cut, universally held definitions for the role and parameters of evaluation is particularly critical and profound. As a relatively new professional field, there is not a wealth of experience upon which to draw. We do not yet have either a sense of "best practices" or a set of easily adapted, conventional techniques for program evaluation that will contribute to the definition of the role and hence to the methods we might adopt in evaluation design.

Adding to the challenge, our projects are typically supported by multiple funding sources, each of which may have its own monitoring requirements and perceptions of evaluation and information feedback needs. As we work to grasp and address the agenda of each of these funding sources, their monitoring and evaluation requirements may be changing as the agencies themselves respond to legislative and other political influences. In short, our ability to create consensus around definitions of evaluation in industrial extension services is tested by the dynamic political environment which surrounds the funding process.

For manufacturing extension partnerships, there is an additional source of potentially differing expectations of the project's evaluation component. By definition, partnerships involve partners, or stakeholders representing a range of backgrounds and orientations. Partners may represent not only political and governmental entities, but such diverse affiliations as public and private academic institutions, service providers, industry sectors and project management. In many cases, the information each needs or expects from the extension service evaluation will be different, resulting in differing or even conflicting definitions being applied to "evaluation" as the partnership strives to develop a consensus-driven evaluation model.

Designing an Evaluation System for the Massachusetts Manufacturing Partnership

The Massachusetts Manufacturing Partnership (MMP) was conceived as an extension service system comprised of five regional centers covering the entire state and coordinated by a statewide management entity. TRP funding, through NIST, was matched through state legislation providing funding to Massachusetts' Executive Office of Economic Affairs. The statewide coordinating function was subsequently contracted to a quasi-public organization. The founding regional partners were identified through a pre-TRP competitive process, and include public and private academic institutions, industry, and industrial advocates and service providers.

Evaluation was identified as a significant program component in the earliest stages of the evolution of MMP, during the development of the TRP proposal. While little was known of prospective federal evaluation requirements, it was clear that our system would pursue evaluation as an integral component of project management. Statements regarding evaluation in the TRP bid described an overall approach that sought to document emerging service delivery models and

assess relative attainment of goals and objectives, while providing ongoing performance feedback to the system, its management and field staff.

With the project funding secured, the first major step in the development of an evaluation plan was to determine the various evaluation needs and agenda of the range of program stakeholders. Governmental evaluation requirements were determined with relative ease -- national workshops and analysis of reporting requirements provided general guidelines for the federal agenda, and performance criteria in the contract for statewide management provided some indication of the state's orientation. But careful attention to the less clearly articulated needs of other program stakeholders would be an important aspect of the challenge of evaluation design.

From the beginning, it was clear that successful implementation of an evaluation plan would depend upon the ability to secure access to reliable data, cooperation in data collection activities and open communication with all parties. To insure cooperation with evaluation efforts, as wide a level of participation as possible was solicited in the evaluation development and design phase, as system-wide "buy-in" to the approach would be greatly facilitated by a "bottom-up" design process. To this end, once we had familiarized ourselves with the federal and state issues, we conducted a series of focus group discussions in each of the five regions in order to better articulate evaluation questions for the system. To the degree that we could generate a consensus-driven evaluation agenda, we felt we would be ensured a high level of cooperation across the system and that the subsequent metrics and data sources would become relatively self-evident.

A Conceptual Framework for Considering the Role of Evaluation

Through the regional discussions we learned that there was no single set of needs or expectations across our system. While the focus was always ostensibly "evaluation", the frames of reference that individuals were applying were widely divergent.

For example, in some sessions discussion focused on NIST's categorization of services for purposes of measuring levels of activity, while in others it examined the means for constructing controlled experiments to test hypotheses as to the relative efficacy of innovative approaches to the restructuring of manufacturing firms, and so on. To a large degree, in fact, the views of evaluation revealed through these discussions tended to be at the extreme ends of the "information needs" continuum suggested earlier. For the most part, expectations focused either on monitoring of project activity primarily to address concerns of accountability (e.g., categorizing and counting firms served and services provided) or on the potential control and manipulation of services towards an ultimate expansion of the state of the art.

Project monitoring typically refers to the collection and use of data for the purposes of basic project documentation and tracking and counting of clients and activities. As such, it addresses fundamental management needs and accountability (reporting) requirements. Project monitoring is typically perceived as a bureaucratic activity and may be attended to routinely and perfunctorily. But while this bureaucratic description paints a picture of an innocuous activity, program monitoring is often a source of anxiety for program employees, and fears of personal appraisal sometimes lead to defensiveness and can interfere with the collection of timely and accurate data. To the degree that monitoring in and of itself is considered to be a program's evaluation, it reinforces the negative stigma often associated with evaluation, as discussed earlier.

Additionally, the perception of evaluation as a perfunctory exercise, and of the data collection associated with it as an end in itself, can result in diminished concern for the activity. When confronted by conflicting demands for one's attention, program staff may deal with monitoring requirements in at least two ways. They may ignore or hastily complete the requirements or they may overemphasize the reporting activity at a cost to service provision. Too often, the experience of program personnel has been that reporting

requirements have been either too burdensome, ultimately unimportant or both.

It is reasonable to assert that while project monitoring is a fundamental activity in program accountability analysis, the data typically gathered to monitor a program do not begin to address the broader evaluation questions inherent to an effort as large as that of a manufacturing extension partnership. These broader questions include: Is the program effective?; Do companies implement actions as a result of services?; Do actions taken by companies result in desired outcomes and subsequent business and economic impacts?; And what services seem to work best for what types of firms with what types of problems?

At the opposite end of the information needs continuum, the objective of experimental (or basic) research methods is the generation of new knowledge through the manipulation of treatment (intervention) conditions and the control of extraneous factors that might otherwise contribute to change. Ultimately, the researcher is interested in establishing causal relationships between interventions and outcomes in ways that lead to generalized findings. In the context of industrial extension, for example, experimental research might be designed to test the relative effectiveness of two alternative approaches to solving a specific industrial need by random assignment of the interventions and control of other variables that would otherwise differentiate the firms.

Like program monitoring, experimental research can offer significant information in support of a program's evaluation. Unfortunately, there is often a limited awareness and communication of basic research findings to program management. While findings often contribute to broader management theory, they are seldom incorporated into the localized evaluation process that supports program refinement. This reflects a general disinterest that many researchers have in applied research and is also symptomatic of common program staff perceptions that experimental research is an unwanted burden whose experimental controls limit their ability to respond effectively and creatively to client needs.

The degree of control over issues such as selection of subjects and allocation of treatment interventions required for systematic experimental research is limited by its impracticality within the framework of a service delivery project. This limitation is clear when one considers the foremost objective of manufacturing extension partnerships: outreach to firms and the provision of badly needed assistance. It must be recognized that a primary differentiation between program evaluation and basic research is the fact that in program evaluation, field constraints and realities will frequently override the methodological concerns that are of primary importance in experimentation.

This is not intended to infer that experimental research is not a credible contributor to the activity of evaluation. As stated earlier, it is particularly useful in the program conceptualization stage, in advance of more localized program refinement activity. In fact, it can contribute to the local refinement process, but evaluation must adapt itself and its methods to the constraints of the program, not the opposite. It is inconceivable, for example, that an extension service might withhold services to a willing and needing client firm to satisfy the need for randomized experimental and control subjects.

Ultimately, while each is a valid response to some of the information needs of a program, in and of themselves neither project monitoring nor controlled research can successfully fulfill what we purport to be the foremost objectives of program evaluation: formative support of program development and management, and summative documentation and assessment of program efficacy, outcomes and impact.

We suggest, therefore, that the most responsive and effective approach to the evaluation of manufacturing extension services rests not in any isolated sphere of activity, but rather in an integrative model of program evaluation, which reflects the objectives of internal and external management and program refinement, project documentation towards potential replicability, and ongoing assessment of the attainment of objectives. To this end, an evaluation must be framed, first and foremost, within the context of the model of change upon which the service delivery system is based.

Through the identification of the components of a model of change, it is possible to identify critical components of program operation, delineate the range of variables associated with each component, determine which among these are feasible to obtain in a reliable manner, and to identify the appropriate source of the data.

Program evaluation as it is found in this model, is the central concern of our information management activities. Information collected in support of the evaluation most certainly feeds the monitoring process and surpasses the value of isolated monitoring activities through periodic, repeated measurement of serviced firms over time. It does so on a range of progress indicators, as a means of assessing potential short and longer term impacts. Constructed in this way, program evaluation provides a strong database across multiple firms which, while not driven by the needs of experimental research, can also support a wide range of quasi-experimental designs, case studies and secondary analyses that exceed the scope of internal program evaluation and begin to address some of the broader concerns and interests of the research community.

Overview of the Massachusetts Manufacturing Partnership's Evaluation Plan

Development of the MMP evaluation plan was guided by the fundamental acknowledgment of evaluation as an integral component of program management. The plan is designed to provide comprehensive documentation of the MMP service delivery model, measure its outcomes and assess its long term impact. It is intended to generate meaningful data that will be analyzed on an ongoing basis to answer practical management questions and assist in the process of program fine-tuning and improvement. Data collected through the evaluation and analyzed in this formative mode will also provide the basis for summative evaluation. That is, over time this data will provide the capability to

assess changes that may be attributed to the services provided by the system to client firms.

The overall goal of the plan is to address the following key evaluation questions:

1. What companies participate in the program? What are their characteristics? In what ways do they differ from other companies that do not participate?
2. What services do companies receive? Do they correspond with the original plan and vision for the system? If not, why?
3. Is the program effective? Do companies implement actions as a result of services? Do actions taken result in benefits for the companies? (i.e., are the services provided to companies leading to desired outcomes and subsequent business and economic impacts?) How have companies benefitted from the services? Which services and/or combinations of services work best for what types of firms with what kinds of needs?
4. Are client firms satisfied with the services they receive? Why? What factors contribute to client satisfaction or dissatisfaction?

Model of Change

At its core, the evaluation plan is based upon a basic model of change. Framed in generic terms, this model simply suggests that in addressing any pre-existing condition warranting change, interventions are prescribed. These interventions seek to achieve specific objectives. It is further suggested that if achieved, these objectives will lead to the attainment of the overall goal, i.e., the desired change.

Placed in terms of the needs of manufacturing firms and of the economy, this change model suggests the following basis for thinking about the mission of an industrial extension service:

- Project managers will provide SERVICES to CLIENTS to address problems that they are facing.
- These services are intended to lead to specific ACTIONS or changes in behavior within

-
- firms. Such actions are taken with the aim of improving internal operations.
- It is expected that companies will see tangible benefits or INTERMEDIATE OUTCOMES from these operations.
 - These intermediate outcomes will lead to BUSINESS IMPACTS, or improvements in the performance of firms.
 - Improvements in business performance are intended to lead to desired ECONOMIC IMPACTS such as increased productivity, increased employment, business retention, etc.

Specific tracking variables and their associated metrics have been identified for each component of the change model. These are summarized in Table 6.1.

The evaluation plan is designed to answer questions related to whether the program has caused companies to take actions that have led to improvements in internal operations, overall business performance and the attainment of desired policy impacts. Through a careful assessment of the relationships among the various components of the model, the evaluation will yield important insights into what works, what doesn't, and why.

Evaluation Design

All firms that receive services through MMP will be included in the system's evaluation component. However, there will be three levels of "intensity" in the evaluation, and depending on a number of criteria, firms will be included at one level of intensity or another. The levels differ in the amount and type of data that will be collected and in the data collection methods that will be employed, with each level building upon the previous.

These three levels of evaluation combine to offer a rich blend of quantitative and qualitative information, but also represent a compromise between methodological considerations and the constraints that surround practical program evaluation. Concern for optimizing the time of MMP project managers, maximizing the appeal of

MMP services to prospective client firms and containing overall evaluation costs were all significant considerations which led us to implement the plan with these three levels of intensity.

At the first level of evaluation intensity, all firms that come into contact with MMP and receive some type of service will be included in Client Tracking. This, supported by client intake and activity logs, will meet the descriptive requirements of the evaluation and provide characterizations of the population served by MMP, services provided and any subsequent actions taken by firms. Other instruments used at this intensity are the Project Closure Form and the Client Feedback Survey which are administered upon completion of a service and offer additional descriptions of services rendered and client firm perception of and satisfaction with MMP services. The Project Closure form will be administered in a repeated measures mode to address the fundamental progress and client valuation measures identified by NIST.

Many client firms will be involved in a second level of evaluation intensity that includes the above, but is supplemented by the use of the Midwest Manufacturing Technology Center's Performance Benchmarking Questionnaire™ in a pre/post, repeated measures design. This pre/post instrument will yield substantive information regarding the effects of MMP projects upon client firms over time, and will provide comparisons with "control" firms (i.e., otherwise similar firms that have not received such services) from within the Performance Benchmarking database. It is intended to aid in the tracking of business and economic impacts that may result from actions taken within firms as a result of MMP service prescriptions.

At the highest level of evaluation intensity, a number of firms will be selected as Qualitative Case Studies. These longitudinal studies, which will involve on-site observation and key-informant interviewing, will allow the careful documentation of MMP interventions and subsequent firm actions and outcomes. In this evaluation component, which relies upon a non-randomized selection process, the value of the detailed descriptive data that will be collected outweighs the cost of the potential selection bias.

Conclusion

The MMP evaluation plan was developed in accordance with the central arguments of this paper and, this being said, may be considered one still-formative model for manufacturing extension partnership evaluation design. These arguments describe the necessity of attending to and building consensus around the clear definition of the meaning and role of "evaluation", and of distinguishing it as one component of a more broadly defined range of information needs that includes everything from rudimentary program monitoring to complex, controlled experimental research. In this distinction we must be clear that the primary objective of evaluation is the support of management through thorough program documentation and the provision of formative feedback and summative evaluation.

This focus is reflected in the design of the MMP evaluation. It clearly attempts to provide a rich blend of quantitative and qualitative information that will document the range of services and service delivery methods employed, and the outcomes of those very specific interventions. It also attends to this task in a manner that is intended to minimize the burden and disruption of evaluation on service deliverers, as demonstrated by the use of differential evaluation intensity levels. It is a valued aspect of the MMP system specifically because it recognizes the primacy of formative and summative needs of program management.

Also implicit in the design of the MMP evaluation are the information needs of other project stakeholders. NIST and other funding agency accountability requirements are recognized and complied with, and parties more concerned with generating controlled experimental research will find valuable information in the rich descriptions and quantitative data that the evaluation compiles. The critical distinction to consider here is that these information needs do not drive the evaluation plan, they are supported through it.

Ultimately, the value of "best management practices" that are shared across the MEP system

will be found in the clarity of the description of services and the accuracy of the subsequent summative assessment of their impact. In this way, evaluation will provide the basis of support for future MEP management and implementation strategies. With so much depending upon the quality of evaluation generated in support of this formative process, it is a recognizable imperative that the obstacles to effective evaluation - the unambiguous definition of its intent, role and function- be deconstructed, and that its complementary role as a provider of data for monitoring and research be understood.

Table 6.1

Variables and Metrics Associated with the Change Model

<u>Components of Model</u>	<u>Variables</u>	<u>Measures</u>
Clients	Industry type	4 digit SIC
	Company size	# employees
	Plant size	# employees
	Sales	Annual sales
	Plant location	ZIP
	Headquarter location	Plant location/other Mass/other non-Mass
	Ownership	Publicly traded (Y/N)
	Ownership	Minority owned (Y/N)
	Ownership	Woman owned (Y/N)
	Union status	Unionized workforce (Y/N)
	Sites	# plants
	Market mix	% sales - engineer to order
	Market mix	% sales - job shop
	Market mix	% sales - repetitive work
	Market mix	% sales - make to stock
	Customer location	% sales in Mass.
	Customer location	% sales other US
	Customer location	% sales out of US (exports)
	Customer type	% sales to final consumer (vs. OEM)
	Defense dependent	% sales defense/military
Production volume	# units in typical order or annual volume	
Production operations	% labor time on assembly	
Product complexity	% shipped jobs w/2 or more parts	
Product value	Average (typical) price of product shipped	
Services	Activity type	NIST classification
	Activity substance	NIST classification
	Provider	Service provider classification
	Delivery mode	Individual or group
	Activity duration	Time elapsed (days)
	Activity level of effort	Person days
	Activity level of effort	Total \$ value of contract (fees)
	Activity intensity	Person days/time elapsed
	Center level of effort	Total person days (all activities)
	Center level of effort	Total \$ value (all activity fees)
	Duration of involvement	Time elapsed: first - most recent activity (days)
	Intensity of involvement	Total person days/duration of involvement
	Customer satisfaction	Satisfaction index
Actions	Redesign existing product(s)	(Y/N)
	Improve existing manufacturing processes	(Y/N)
	Purchase production equipment	(Y/N)
	Adopt EDI	Use (Y/N)
	Adopt costing system	Incorporate machine use (Y/N)
	Adopt costing system	Incorporate floor space for inventory (Y/N)
	Adopt costing system	Incorporate downtime (Y/N)
	Adopt CAD/CAE	Use (Y/N)
	Adopt CAD-CAM	Use (Y/N)
	Adopt CNC	Use (Y/N)
	Adopt software for: sched., inv. control or purchasing	Use (Y/N)
	Increase computer usage	# keyboards/keypads in use (non-office)/shop floor employee
	Increase computer usage	% employees using comp. at least once/week
	Change plant layout	(Y/N)
	Adopt energy saving practices	(Y/N)
Adopt waste reduction practices	(Y/N)	

Table 6.1 (continued)

Variables and Metrics Associated with the Change Model

<u>Components of Model</u>	<u>Variables</u>	<u>Measures</u>
Actions	Train workforce	Training \$/employee
	Train workforce	Training \$/shop floor employee
	Adopt statistical Q.A. program	(Y/N)
	Registered to ISO 9000	(Y/N)
	Adopt production teams	% shop employees in production teams
	OTHERS	(as identified by field staff)
Intermediate Outcomes	New product introduction	% sales for products introduced w/i past 3 yrs
	Setup time	Change in # hours/typical setup
	Mfg. lead time	Change in # days prod. start-end
	On-time delivery	Change in % deliveries made on time
	Expedited jobs	Change in expedited jobs/total jobs
	Scrap rates or yield loss	Change in scrap rate or yield loss
	Customer rejects	Change in % jobs or lots rejected by customers
	Energy costs	Change in total energy costs
	Waste disposal costs	Change in total waste disposal costs
	Inventory turns	Change in sales/inventory
Labor productivity	Change in value added/employee	
Business Impacts	Sales	Change in annual sales
	Exports	Change in exports/sales
	Profitability	Change in (sales-cogs)/sales
	New customers	Change in % sales to new customers
Economic Impacts	Employment	Change in # employees
	Employment	Change in # shop employees
	Average wages	Change in total payroll/# employees
	Average shop floor wage	Change in shop floor payroll/# shop employees
	Labor productivity	Change in value added/employee
	Investment in hardware/software	Actual dollar investment
	Investment in capital equipment/ buildings	Actual dollar investment