Working Hard
Working WELL &

Performance Monitoring and Evaluation:
A Complimentary Relationship

by David E. K. Hunter

———

An excerpt from the book *Working Hard & Working Well*, which is
available in full at leapofreason.org/workingwellbook

© 2013 by Hunter Consulting, LLC

Performance Monitoring and Evaluation: A Complementary Relationship

Performance monitoring often is confused with evaluation. While they are certainly related, there are real differences between the two. As discussed above in connection with Figure 1, both kinds of knowledge production are essential for high performance management.[63]

Simply put, performance monitoring is necessarily a forward-looking activity. It asks, "What are we accomplishing and how do we need to improve?" As an organization monitors and learns from its work, it will adjust what it measures and tracks in order to do a better job of managing its performance, strategically and tactically. Thus measurement is an iterative process, and its use and methods are viewed entirely within the context of the organization's work and its ability to drive performance toward the achievement of targeted results.

---

63. Developed in collaboration with Steffen Bohni Nielsen.

Table 6. Comparison of the Complementary Activities of Performance Monitoring and Evaluation

DATA-RELATED ACTIVITY	ITEM CHARACTERISTICS	PERFORMANCE MONITORING	EVALUATION BOTH FORMATIVE AND SUMMATIVE
Program planning, implementation, and management	Purpose	Evolving; iterative	Negotiated up front
	Focus	Broad; ongoing	Issue-specific; retrospective
	Timing	Constant/ongoing	Periodic
	Types of data collected	Inputs, outputs, outcomes	Inputs, outputs, outcomes, impacts
Obtaining data	Collector	Internal staff	External staff
	Collection intervals	Ongoing, "real time"—forward-looking	Rare—at large intervals (monthly/quarterly/yearly)—backward-looking
	Data validity	Presumed	Must be tested as part of the research
	Collection tools	Internal data (client demographic, staff efforts, client service utilization, outcomes), assessments, surveys, etc.	Desk research, public databases, interviews, observations, surveys, mining of information systems, new measurements using validated tools, etc.
Using data	Effectiveness	Causality of outcomes presumed for management purposes	Causality is a core issue that must be established scientifically; often a contribution analysis is more meaningful than an attribution analysis
	Organizational learning	"Real time" or ongoing	Infrequent—ex post facto
	Tactical utility	High ("real time" flow of data)	Low (very slow flow of data, often "stale")
	Strategic utility	Low to moderate	High
	Number of users	Many (all of the organization's leaders, managers, staff)	Few (mostly the organization's leaders and managers)

On the Validity of Performance Data

For the purposes of performance management, it is necessary to assume the validity of performance data. Hence it is essential for organizations to monitor the integrity (timeliness, accuracy, and completeness) of their performance data and eventually to test the validity through a formative evaluation (as discussed above).

In contrast, evaluation is backward-looking. It asks, "What has the organization done and accomplished? Does this meet our goals?" It provides the external reference points—including the use of validated measurement tools—to assess how well the organization's performance data capture reality (how valid they are). Measurements are designed up front, at the beginning of an evaluation, and then should be held constant for its duration (although admittedly there are many examples of evaluations that do not meet this standard—and hence are of questionable worth). Failure to hold them constant removes the high value of mapping performance data against a stable reference framework. (Imagine measuring your child's growth using an elastic tape measure where the length of inches and feet shift over time.)

Evaluation data are always *ex post facto* and "stale"—and thus their usefulness for tactical performance management is low. But they are extremely useful for holding an agency's performance data, especially regarding intermediate and long-term outcomes, up to the light of rigorous scrutiny—a function that gives such data great strategic utility.

With regard to tracking outcomes, it is important to highlight another way in which performance management and evaluation are complementary: **For performance-management purposes, it is essential to focus on short-term outcomes and the incremental progress that clients make toward them. In contrast, evaluations usually focus on intermediate and long-term outcomes because they are the measures of the value that an agency has produced.** Here short-term outcomes are seen as key inputs into the achievement of more enduring intermediate and ultimate outcomes.

It is worth briefly noting that there is a general sequence of evaluation activities that an organization should follow.

A Useful Sequence of Evaluation Activities

Step 1. Track performance data for several years until it is clear that the organization is enrolling the people it intends to help (target population) in its core programming, delivering the services in the ways and at the dosage levels it has specified, and achieving the outcomes it has targeted.

Step 2. Undertake an assessment of the reliability of the organization's performance-management system (data and the processes for using it to make adjustments in order to drive the achievement of results). Review the findings and make the organizational adjustments that are indicated. Take at least a year to test how well the organization has implemented these adjustments. Then, if they have been sustained and performance is high, it would be reasonable to move to the next step.

Step 3. Undertake an external "formative evaluation" that tests the validity of the agency's performance data regarding program enrollees, staff competencies, delivery of specified services according to implementation standards, and service recipients' achievement of short-term and intermediate (and, where the data allow, long-term) outcomes.

Step 4. Take the time to absorb the lessons of the formative evaluation and develop whatever plans are necessary to improve the areas of performance that were shown to be less than optimal (there are always some). This should identify specific organizational competencies and capacities, systems, and processes that will need adjusting to improve performance.

Step 5. Undertake a rigorous assessment to see how well the targeted organizational elements have been upgraded or advanced and whether the associated performance improvements have been achieved.

Step 6. Run the organization for at least another two years while sustaining high performance and program effectiveness as measured by service recipient outcomes. Repeat Step 5. If the organization has fallen short in any areas, make adjustments and then repeat Step 6.

Step 7. Once the organization has been running at high levels of quality and effectiveness as specified in its implementation standards—in other words, it is meeting its performance standards as codified in its objectives—and has been able to do so for several years, the time may well be right for a "summative evaluation" that, among other things, tests the impact that the organization's services have had on the lives of participants. However, if the organization's scale of programming is too small for a rigorous evaluation to generate statistically significant data, the next step will be to grow the programming capacity to a level that supports a scientific evaluation approach—and then to repeat Steps 1 through 6 before moving on to a summative evaluation.

Unfortunately, few organizations and few funders are willing to exercise such discipline. Without it, evaluations will continue to be rigorous but useless—well measured but spuriously precise in their findings. The efficacy of programs and services to help participants gain targeted outcomes will be found wanting, but the fault may lie not in a weak model or approach but rather in the delivery of services by an organization lacking the capacity to manage performance strategically and tactically.

This, needless to say, is tragic. **It undercuts our ability to learn what works to help various populations improve their lives and prospects, and it gives us wrong ideas about what doesn't work.** Most tragic is that unless an evaluation[64] specifically looks at performance management, it will fail to identify organizational capacities and competencies that need to be developed in order to deliver programming reliably, efficiently, effectively, and at high levels of quality!

---

64. This is one way in which a formative evaluation can, depending on where an organization is in its development, be more useful than an impact evaluation. The former looks at key aspects of organizational functioning as well as the specific contributions of elements of programming to the production of outcomes; pure impact evaluations treat organizations simply as context, and tend not even to look inside programs to see how their elements work—hence the term "black box evaluation" that is often applied to such studies.