

# Performance Measurement and Evaluation

October 2018

Bringing evidence to bear in decision-making is a critical component of effective and efficient government. **Performance measurement and evaluation** are two key tools available to help policymakers and program managers develop systematic evidence, understand how well policies and programs are working, and identify possible improvements. Both evaluation and performance measurement generate information that falls along the continuum of evidence, serve as methods for systematic assessment, and aim to facilitate learning about and improve results of government activities. At the same time, there are important differences between these methods that dictate what each can tell us about programs and policies.

## Performance Measurement

## Evaluation

### DEFINITION

- Ongoing collection, monitoring, reviewing, and reporting of data on pre-selected measures related to level and type of activities, products and services delivered, and outcomes of activities
- Individual, systematic studies to examine how well all or part of a program, intervention, policy, regulation, or other government activity is working

### PURPOSE

- Measuring progress toward pre-established goals and targets
- Determining whether an activity is achieving its stated output/outcome objectives and making adjustments if it is not
- Serving as an early alert system in the case of significant changes in operations
- Assessing the effectiveness of a program, intervention, policy, or regulation, compared with its absence or with one or more alternative approaches
- Establishing a causal relationship between an activity and the outcomes experienced by those affected by it
- Addressing questions about implementation, variations in effectiveness across different settings or populations, and contextual factors

### DATA AND ANALYSIS

- Data is largely quantitative
- Data points assessed against targets or compared to previous data for same measure, in order to detect trends over time
- Data and analytical techniques are guided by the evaluation questions
- Generally includes both quantitative and qualitative data
- In the case of causal studies, requires complex methods to isolate impacts from other influences

### RESPONSIBLE PARTY

- Usually an internal function undertaken and managed by the staff of an agency/funder as part of its routine operations
- Often carried out by independent researchers who are external to the agency/funder to ensure independence and impartiality
- Requires technical expertise in advanced methods

### EXAMPLES OF QUESTIONS ADDRESSED

- Did the program meet its stated output goals?
- How many individuals participated?
- What percentage of people who participated in a program reached a certain goal (e.g., got a job, completed college)?
- In each program site, what was the average length of time it took participants to complete a program?
- Why did certain individuals engage or not engage in a program?
- How many people reached a certain goal (e.g., got a job, completed college) as a result of access to a program, compared to those who did not have access?
- How does the implementation of a program differ across sites, and how do those differences affect participants' experiences?

## HOW CAN PERFORMANCE MEASUREMENT AND EVALUATION WORK TOGETHER?

While often undertaken separately, collaboration between performance measurement and evaluation teams can lead to stronger evidence-building. Ways the two can work hand in hand include:

- Performance measurement can help identify priority questions to be addressed by evaluations, informing decisions about allocating evaluation resources.
- Evaluation findings can clarify what indicators are predictive of an activity's success and should be tracked in performance measurement.
- Evaluation can provide context and potential explanations for variation over time or across sites revealed by performance measurement.
- When performance measures suggest that many participants in a program experience a certain outcome, evaluation can confirm (or refute) whether that is directly attributable to the program by comparing outcomes seen in a control or comparison group when possible.
- Performance measurement can suggest to evaluators what types of indicators are important to program operators and might thus be useful to include in selecting evaluation measures.

### CASE STUDY #1

A government agency that administers a large formula grant program to states looked at performance data and saw that they were falling short of their enrollment targets. Staff observed that a significant portion of individuals who were eligible for the services funded by the grants were not receiving them. This conclusion, drawn from the performance data, motivated the agency to implement a behavioral science-informed intervention aimed at “nudging” participants to take advantage of these services. The program ran a randomized controlled trial evaluation of this intervention in order to determine whether it did in fact increase uptake of services as intended, compared to service uptake without the intervention. The main outcome of interest in that study was the same performance metric: the number of individuals who participated after receiving the behavioral “nudge” compared to the number of individuals who participated without having received the intervention. Performance measurement processes inspired an evaluation that was ultimately aimed at finding ways to improve upon a particular performance metric that was important to the program. Simultaneously, there is an ongoing impact evaluation of the overall program that looks at whether individuals who received these services experienced better outcomes than a control group of individuals who did not.

### CASE STUDY #2

A multi-site national program had been tracking performance for over a decade, collecting data on various measures and comparing it to goals for each measure. The performance information was used for a range of purposes, including to reward sites, pay incentive bonuses to staff, and decide whether to renew existing site contracts. When the program underwent a large-scale random assignment evaluation, researchers saw an opportunity to compare the performance data with impact evaluation data by analyzing whether participants at sites that consistently met performance targets were likelier to experience better outcomes than a carefully selected control group that did not participate in the program. This independent study revealed that there was a weak connection between how sites were doing on the performance measures and the extent to which their participants were faring better than the control group. Sites that appeared to be top performers based on their performance data did not always have the biggest impacts on participants, and sites that had reported lower performance did not necessarily have less of an impact on participants' outcomes. The research was additionally able to use data to identify some possible causes for this lack of connection, such as the fact that the higher-performing sites were on average serving higher-ability participants from the outset. This instance demonstrates how evaluation can serve as a crucial supplement for performance data.