Performance Measurement & Program Evaluation:
A Suite of Evaluative Insights

One of the most frequently asked questions among new and even seasoned evaluators is "How are performance measurement and program evaluation related?" In this brief, we provide insights about the usefulness of these two inquiry methods and describe how they complement one another.

Origins and Definitions

Performance measurement

- Is defined as "the ongoing monitoring and reporting of a program's accomplishments and progress, particularly towards its pre-established goals" (GAO, 2021, p.3).

- Has origins in public administration as part of the New Public Management movement in the 1990s (McDavid, Huse, Hawthorn, 2019).

- Answers the general line of questioning: "**What** occurred or what is occurring?"

- Was developed to inform **performance management**. By interpreting patterns in performance measures, it was envisioned that program managers could adjust the existing strategy, if needed, thereby improving the likelihood of accomplishing programmatic objectives and goals. Potential adjustments include implementing strategic planning, budgeting, and personnel management (Poister, Aristigueta, & Hall, 2015).

- Includes the calculation of key indicators from qualitative or quantitative data. Key performance indicators, or KPIs, are calculated and reported with regular frequency—weekly, monthly, quarterly, bi-annually, annually—depending on such things as information need and data availability. To interpret results, the KPI is often compared to something else, such as a pre-determined target or benchmark, another program's performance, or past performance for the same program.
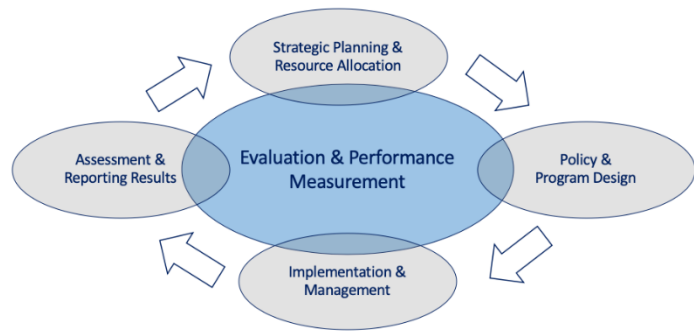
Program evaluation

- Has been defined in several ways. We adopt Patton's definition (2008), given its emphasis on the systematic nature of evaluation and making use of evaluative findings: "the systematic collection of information about the activities, characteristics, and results of programs to make judgments about the program, improve or further develop program effectiveness, inform decisions about future programming, or increase understanding" (p. 39).

- Has multiple roots (Alkin, 2013) but primarily originates from the social sciences during the 1960s. Program evaluation was a response to calls for increased accountability for public spending in the War on Poverty (Shadish, Cook, & Leviton, 1999).

- Answers the general line of questioning: "*Why* did it happen?" and "*How* did it happen?"

Complementarities

Performance measurement and program evaluation can both inform decision making. However, these evaluative inquiry methods answer different questions. Performance measurement can provide high-level insights about what is occurring or how much of something is taking place. Program evaluation dives deeper into the quality of implementation and the degree and value of change, often by applying social

science research methodologies (Rossi, Lipsey, & Henry, 2019). Though they often compete for the same resources (Nielsen & Hunter, 2013), the two methods are complementary forms of evaluative inquiry and should not be used as substitutes for each other (McDavid et al., 2019).

Performance measurement and program evaluation can interact to create a robust and comprehensive guidance system using these, and other, strategies:

The Performance Management Cycle from
McDavid et al. (2019), Figure 1.1., p. 9

- Using tools of the program evaluation trade, such as stakeholder identification and engagement techniques, logic modeling, and measurement principles, to develop robust performance measures.

- Taking observations from performance measurement data to stimulate questions to examine through program evaluation.

- Using performance measurement data to answer program evaluation questions.

- Monitoring unintended positive or negative effects of a program identified in an evaluation by creating new outcome indicators and integrating them into ongoing tracking.

- Building the capacity to effectively engage in evaluation and make use of the insights it produces. Incorporating program evaluation and performance measurement into an organization's routine operations can advance its capacity to commission, conduct, and use evaluation (Bourgeois, 2016). These routine operations can increase the extent to which stakeholders value evaluative insights and have realistic expectations about the efforts these processes require.

Is the time right for performance measurement or program evaluation?

- You might benefit from performance measurement if time is of the essence and there is a regular, perhaps frequent, need for general insights about how much of something is happening, when, and for whom. These insights may be useful for reporting to funders or for making minor programmatic adjustments.

- Performance measurement can produce data that may be needed in future evaluations. If you are not ready to develop an evaluation study about programmatic outcomes (e.g., due to a program's current level of maturity), you may want to consider establishing regular, ongoing data collection about outcomes *now*. Having a system in place to collect high-quality outcome data may enhance your ability to perform successful outcome evaluations in the future.

- It is likely time to consider engaging in program evaluation when many questions arise while examining performance measurement patterns, and further analyses of the data are not able to answer these questions.

- When priority evaluation questions relate to *causation*—in particular, the extent to which the program, intervention, or other activity results in, or contributes to, changing outcomes—it is time for program evaluation. Performance measurement can address questions related to correlation but typically falls short in achieving the internal validity required to confidently examine causality.

Note: Several resources explore the similarities and relationship between performance measurement and program evaluation. (See reference list.) We consulted the following sources extensively in developing this brief: Bourgeois (2016), McDavid et al. (2019), and Nielsen and Hunter (2013).

References

Alkin, M. C. (2013). *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed.). Sage.

Bourgeois, I. (2016). Performance measurement as a precursor to organizational evaluation capacity building. *Evaluation Journal of Australasia*, *16*(1), 11–18.

McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2019). *Program Evaluation and Performance Measurement: An Introduction to Practice* (3rd ed.). Sage.

Nielsen, S. B. & Hunter, D. E. K. (Eds). (2013). Performance management and evaluation. *New Directions for Evaluation*, *137*, 1–123.

Patton, M. Q. (2008). *Utilization-Focused Evaluation* (4th ed.). Sage.

Poister, T. H., Aristigueta, M. P., & Hall, J. L. (2015). *Managing and measuring performance in public and nonprofit organizations: An integrated approach* (2nd ed.). Jossey-Bass.

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation: A systematic approach* (8th ed.). Sage.

Shadish, W. R., Cook, T. D., & Leviton, L. C. L. (1991). *Foundations of program evaluation*. Sage.

U.S. Government Accountability Office. (2021). *Program evaluation: Key terms and concepts [GAO–21–404SP]*. https://www.gao.gov/assets/gao-21-404sp.pdf.