MATHEMATICA
Policy Research

IssueBRIEF

Jonathan Ladinsky

Developing an Effective Performance Management System: Lessons for the Implementation of WIOA

The Workforce Innovation and Opportunity Act (WIOA) of 2014 helps job seekers access education, training, and support services to succeed in the labor market and helps match employers with the skilled workers they need to compete in the global economy. To assess states' performance, the legislation modifies Workforce Investment Act measures so that states and other stakeholders can use the federal system to assess performance and hold programs accountable.

This issue brief provides guidance on what it will take to develop a first-class performance management system that allows states to collect and report accurate data. It provides recommendations to help federal, state, and local agencies ensure that their performance management systems support accountability and can help them manage their programs. It also provides guidance to help federal and state labor and education agencies develop and implement the new systems to generate quality performance data at a reasonable cost.

In preparing this issue brief, we draw on lessons learned from the implementation of performance management systems for other programs, including the Workforce Investment Act of 1998, the Department of Labor's common measures, Temporary Assistance for Needy Families, and several discretionary and nondiscretionary Department of Education programs.

WIOA integrates the U.S. Department of Education's Adult Education and Family Literacy program and the Vocational Rehabilitation program into the federal labor accountability system, which includes the U.S. Department of Labor's Workforce Investment Act (WIA) adult, dislocated, and youth programs; and the Wagner-Peyser Employment Service program.

WHY PERFORMANCE MANAGEMENT SYSTEMS MATTER

Performance management systems can serve many purposes, the most significant of which are accountability and program management. Policymakers, program managers, the public, and other stakeholders use measures to assess the performance of state and federal programs and to compare performance across time and states when making program and funding decisions. Generally, performance measures used for accountability measure macro-level events, such as the entered employment rate.

Performance management systems also help staff manage programs. The data used to calculate accountability measures, plus additional data and performance measures that states and local offices use to make short-term decisions, can help improve program outcomes. These data and measures focus on more micro-level events, such as details of case management and program services.

Successful performance management systems must provide information that is accurate, complete, comparable among states and across time, standardized, stable, transparent, and usable for managing and assessing a program's success. These information attributes often overlap. For example, only accurate performance information can be useful, and only stable information can be comparable.

WHAT WE KNOW

For the past 15 years, Mathematica has helped analyze, develop, and manage performance management initiatives for clients in education, health, labor, and welfare. This work has provided us with lessons that can help federal, state, and local program staff develop and implement a WIOA performance management system. Several steps are necessary to develop an effective system.

that including these participants would complicate the measure and make it difficult to assess true program performance. Federal agencies need to think through these details and should do so after receiving input from important program stakeholders.

Experience suggests that developing measures that are easy to understand is important. Doing so helps states and local offices more effectively collect and report the necessary data and helps policymakers and other stakeholders better assess performance. It is important to differentiate between measures that are difficult to calculate and measures that are difficult to understand. For example, the employment-related measures in the Adult Education and Family Literacy program allow states to survey a sample of adults who have left the program. Although determining the sample size needed to achieve a precise estimate can be a difficult

8 steps in creating a first-class performance management system

1. Develop measures that are easy to understand and calculate.
2. Build a robust, complete set of measures.
3. Create detailed technical specifications for the measures and data.
4. Fully develop the system before implementation.
5. Create documentation standards and arrange for required data concurrently with system implementation.
6. Provide states with a web-based reporting and validation application.
7. Provide training and ongoing technical assistance to states.
8. Collect additional data to analyze program management and evaluation.

① Develop measures that are easy to understand and calculate.

Authorizing legislation for government programs often identifies many of the performance measures required for program accountability but rarely provides the details needed to calculate these measures. For example, the WIA legislation required states to report on participant entry into unsubsidized employment, but it did not stipulate excluding from this measure those participants who were already employed when they enter the program. The Department of Labor (DOL) added this exclusion after deciding

statistical problem, doing so allows the results to be presented so that stakeholders can easily understand them.

Measures that are complicated to calculate can also be difficult to specify and program. As a result, they may yield results that are inaccurate or difficult to compare across states. When complicated specifications are needed, programs benefit from having a federally developed and supported online information system that calculates performance measures for states, local agencies, and grantees. For example, the WIA skill attainment

rate for younger youth measure was complicated because it had many conditions and was not specified. The measure required out-of-school, basic-skills-deficient youth to set at least one skill attainment goal per year and to attain it within a year of setting it. Complicating the measure was the fact that participants could set as many as they wanted, and only the first three goals were included in the performance calculation. Because the measure had so many conditions and no clear federal specifications, states had to create their own definitions and specifications, which varied across states, thus reducing the comparability of their performance results.

The WIOA skill gains measure could be similarly complicated. It is meant to assess skill gains for those in postsecondary training, like the common measures' literacy/numeracy rate, which requires states to test youth to determine if their skills improved within a year. Differences in determining how to measure skill gains, the changing definition of skill gains, and the complicated logic that affected whether and when the measure applied to a specific youth all make it difficult for states to understand and calculate. For these types of measures to be accurate, comparable, and transparent, developers should be sure to define and specify them. In addition, they should consider how the measures will be used, as well as the cost for states to collect the data to calculate and report accurate results.

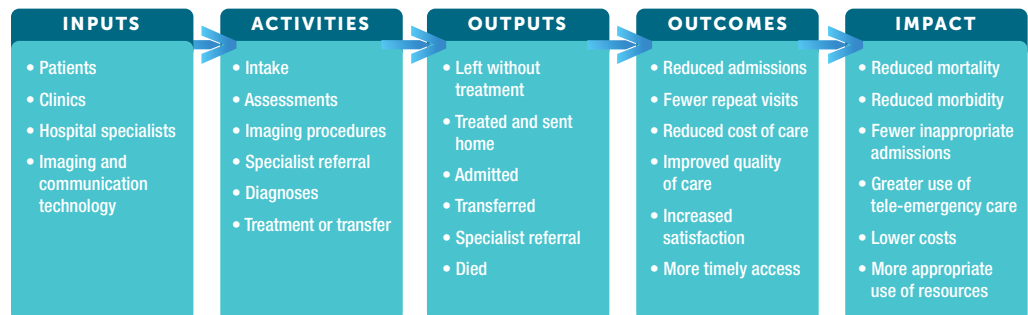### ② Build a robust, complete set of performance measures.

No performance management system covers all aspects of a program—they all have gaps. Avoiding gaps for a program's most important aspects—and being aware of the consequences of existing gaps—are key in developing a strong management system. A logic model can help identify important program activities and outcomes and show how a program achieves its short- and long-term goals.

Logic models, and the thought process involved in creating them, help identify what aspects of a program to assess; however, they do not provide an integrated view of performance measures.

Decision matrices are tables that identify the key performance drivers (that is, the data elements used to calculate performance measures or categories identified in the logic model) and characteristics of the program and its participants, show how these performance drivers and characteristics can vary, and indicate how they relate to performance outcomes. The final table contains all relevant combinations of the performance drivers and program characteristics used to calculate all the measures, or related groups of measures, in the system. For example, a table for the common measures' entered employment rate (EER) and employment retention rate would review all possible combinations of the data elements for these measures—including employment status at participation and in the first, second, and third quarters after the exit quarter. Policymakers can use these tables to analyze gaps in the accountability system and decide whether to close them.

The graphic on page four is a simplified sample of a decision matrix for the entered employment and employment retention rates. Columns B through E identify the key variables that determine the outcomes. Columns F to I identify whether each combination is included in the measures and whether it is a success (included in both the numerator and denominator), or excluded from the measure (included in neither the numerator nor denominator). Looking at the combinations, you can see that those participants who match the second category are not in either measure. Other categories can be added to the matrix if policymakers and program managers consider them important. For example, disability status could be incorporated by adding a disability status column to the table and then creating the necessary combinations for all the variables in the table, including disability status.

Logic model to guide selection of measures for tele-emergency care

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES	IMPACT
• Patients • Clinics • Hospital specialists • Imaging and communication technology	• Intake • Assessments • Imaging procedures • Specialist referral • Diagnoses • Treatment or transfer	• Left without treatment • Treated and sent home • Admitted • Transferred • Specialist referral • Died	• Reduced admissions • Fewer repeat visits • Reduced cost of care • Improved quality of care • Increased satisfaction • More timely access	• Reduced mortality • Reduced morbidity • Fewer inappropriate admissions • Greater use of tele-emergency care • Lower costs • More appropriate use of resources

Decision matrix for entered employment and employment retention rates

A. #	B. Employed at participation	C. Employed in 1st quarter after exit	D. Employed in 2nd quarter after exit	E. Employed in 3rd quarter after exit	F. Entered employment numerator	G. Entered employment denominator	H. Employment retention numerator	I. Employment retention denominator
1	Yes	Yes	Yes	Yes	No	No	Yes	Yes
2	Yes	No	N/A	N/A	No	No	No	No
3	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4	No	No	N/A	N/A	No	Yes	No	No

States and local offices can also use logic and decision matrices to develop their account-ability and performance management systems. Even when federal agencies do not incorporate program activities or outcomes into their performance management system, states and localities may decide that these activities and outcomes are important for assessing and managing their programs. For example, Temporary Assistance for Needy Families seeks to promote job preparation and work, among other goals. The program office requires states to report their work participation rate, which is calculated using the hours that participants spent doing each of 12 work-related activities. However, it does not assess other activities and outcomes such as substance abuse treatment, which can impact employment-related outcomes for some participants. Although Congress may have intentionally excluded substance abuse treatment from the federal accountability system, states and local offices will have clients with this condition and may choose to include it in their system.

③ Create detailed technical specifications for the measures and data. Performance measure specifications should be unambiguous, detailed, technical instructions on how to calculate a measure. Specifications translate the policy behind the measures into technical instructions for programmers, who are not experts in policy. They do this by identifying the data elements used to calculate the measure, the relationships among these elements, and the values they must take to be included in the measure (see box). As a result, they ensure that measures are accurate, transparent, and standardized.

Unfortunately, program staff are often in a rush to implement a required system and may not develop these specifications, at least initially.

The combination of time pressure and vaguely defined measures often limits the program office's ability to provide technical specifications. Agencies and programs need to refine the broad measures in the legislation to help states understand how to calculate them.

Data element specifications are essential for creating accurate performance measure specifications. They define the properties of the data, including the data name, valid values, definitions, and appropriate edits. The definitions clarify what the data element represents, sometimes by citing the legislation or regulations. Edits address potential data quality problems, which can be simple issues, such as valid values, as well as complicated ones, such as duplicate record detection.

The box provides the specification for the entered employment rate numerator, which makes it easier for programmers to calculate and analyze results. The data elements come from the WIA Standardized Record layout. By identifying the variables, their values, and their relationships, the specification makes it easy for programmers to understand the formula used to calculate the measure.

> **Performance measure specifications** should be unambiguous, detailed, technical instructions on how to calculate a measure.

> **Technical specification for the adult entered employment rate numerator**
>
> Numerator. Count of records where DATE OF EXIT >= Start of the report period and DATE OF EXIT =< End of the report period and EMPLOYMENT STATUS AT PARTICIPATION = 2 or 3 and EMPLOYED 1st QUARTER AFTER THE EXIT QUARTER = 1.

④ Fully develop the system before implementation. Implementing a system quickly and adjusting it afterwards may seem attractive to program offices, but modifying it after launch can be costly, can hurt the system's usability, and can reduce users' trust in the program. Planning and defining the system is important, and it takes time. When federal agencies require states to implement a system, states expend significant IT and staff resources. If federal agencies modify the system after implementation, states may need to reprogram and retest case management and other IT systems; train staff and case managers to implement new requirements; and modify vendor contracts, which all cost time and money. Furthermore, old performance data may no longer be useful for program management, which means that performance accountability, analysis, and results will not be comparable over time.

The implementation of the literacy/numeracy rate illustrates these problems. This completely new measure for DOL employment programs required states to collect two years of data before they were able calculate complete, accurate results. However, DOL could not wait two years for results and required states to report on the measure when it was implemented in 2005. This reporting deadline did not provide sufficient time to develop clear performance and data element specifications. As a result, states received multiple updates to the specifications and data require-ments in the first three years of implementation, which cost them time and money.

⑤ Create documentation standards and arrange for required data concurrently with system implementation. Performance management systems need high quality data to meet their goals and ensure accurate results. Identifying required eligibility and performance data and the documents or systems from which to collect these data is the first step in this process. Program offices should create documentation standards before implementation, because doing so afterwards requires rewriting of procedures and retraining of staff, which can be expensive. In addition, some data will be under the control of other agencies and systems, and it will take time to develop data sharing agreements and processes to acquire the data.

WIA identified data elements and required sources, but only after states started to collect and report performance data. It took more than a year for states to implement the data collection and documentation requirements.

Like WIA, WIOA programs will need unemployment insurance (UI) wage records to calculate the employment and earnings measures. Table 1 looks at how WIA, DOL, and states addressed some of the issues WIOA will confront in obtaining wage records.

Table 1. How WIA addressed challenges WIOA may face in obtaining wage records

Issues for WIOA	Approaches taken by WIA
1. State workforce programs need access to wage records for participants in their programs, some of whom may find employment in neighboring states.	1. States negotiated agreements with UI offices of their state and neighboring states. In some instances, groups of states created regional agreements so that each had access to the necessary wage records for all participants in the state.
2. Participants can commute or move to other states for work, beyond neighboring states, so programs will likely need access to wage records from multiple states.	2. DOL created the Wage Record Interchange System (WRIS), which provided member states with access to participant wage records from all member states.
3. State UI offices and regulations limit access to wage records, which is problematic for grant-funded programs and non-DOL programs that are normally unable to gain access to these wage records.	3. DOL created the Common Information Reporting System (CRIS) to help grant-funded programs use wage records to calculate performance measures. Programs submitted social security numbers to a WRIS member state, which then used the WRIS to obtain the necessary wage records, calculate the employment and earnings measures using them, and provide each grantee with aggregate performance outcomes, rather than individual wage records.
4. UI wage records do not cover some types of employment, such as agricultural labor.	4. DOL allowed the use of supplemental (non-wage-record) sources for determining employment and retention, but not earnings.

Whatever the approach, UI wage records are required for WOIA. DOL programs already have access to these records. State and federal Vocational Education and Adult Education and Family Literacy programs will need access to these data, and they should arrange for it sooner rather than later.

6 Provide states with a web-based reporting and validation application. Providing a centrally developed web-based application to calculate performance measures and help states improve their data quality can be a cost-effective way to improve the accuracy and comparability of results. WIOA standardizes many federal reporting and data requirements for the core programs, so each program reports the same performance measures. The federal government could develop one web-based reporting and validation tool for all 50 states, as well as the District of Columbia and Puerto Rico, and all core programs, rather than having each state and program develop its own. A federally provided application would improve comparability and accuracy because it would use the same formulas to calculate each state's and program's WIOA performance measures. This application would not need to be program-specific; therefore, it would not calculate other reporting requirements that individual programs might have. Furthermore, states would not need to submit aggregate performance results. They would only need to submit individual, de-identified participant records in a standardized format that applied to all core programs. The application would use these records to calculate performance and provide analytical functionality to help states understand their performance outcomes.

Under WIA, states used federally developed software to calculate performance for their annual reports and assess the quality of these reports. This process improved the accuracy of states' reports and ensured they were comparable to reports from all other states using the software. Initially, WIA used distributed software—meaning each state had to install it on its own computer system. Because states configured their systems differently, DOL invested significant time and money helping states install the new software. When updates were issued, states sometimes had to uninstall the old software and install the new. With each update, states ran into new problems, which

created tensions between DOL and the state WIA office. Web-based software solves these problems—there are no installation issues, and users do not have to do anything when the system is updated. DOL has benefited from web-based software with the Senior Community Services Employment Program (SCSEP) Performance and Results Quarterly Reporting system (SPARQ), and more recently with WIA.

In addition to its reporting and data quality functions, SPARQ provides SCSEP grantees with case management and analytical functions. A case management system would not work for the core WIOA programs, however. In most states, these programs are integrated with a wide range of other state and federal programs that a WIOA case management system would not support. Furthermore, states require more information, some of it state-specific, to manage and improve their programs than a federal case management system would provide.

7 Provide training and ongoing technical assistance to states. No matter how well designed a performance management system is, states will need help understanding and implementing it. With WIA and the common measures, DOL provided training to help states and local offices hear from and question policy and technical experts about changes to the system and schedule. Most of these trainings were regional, although some of the more technical discussions have been webinars. During these trainings, states hear about the policy and reporting requirements of the new program. They are also able to ask questions about issues associated with implementation, including specifications and data requirements. Throughout the data collection and reporting process, DOL provided technical assistance to states to answer questions about their specific issues with the measures and helped to diagnose performance result and data quality problems.

8 Collect additional data to analyze program management and evaluation. In this era of big data and analytics, state and local program offices need to take advantage of the wealth of data they collect to improve management and provide better service to participants. However, federal programs generally develop systems for federal accountability, not program management, and not for states and local offices to

manage their day-to-day operations. Federal systems focus on a narrow set of measures and the data needed to calculate them, which can be quite limited and collected months after a participant leaves the program. By the time states have reported their data to the federal agencies, the policies and measures they assess may have changed. For example, WIA required states to report performance for participants up to 1.5 years after they exited the program. WIOA's two employment measures will also assess performance for participants long after they leave the program, up to 1.75 years after they exit. These data serve the federal program offices' needs, but they do not support state and local office accountability and management.

Developing more robust performance management systems can help states and local offices better manage their programs and assess their performance in less time, while also incorporating federal performance requirements.

Developing more robust performance management systems can help states and local offices better manage their programs and assess their performance in less time, while also incorporating federal performance requirements. Doing so requires a better understanding of the factors that lead to successful program management and outcomes. Based on this understanding, state and local offices can develop a system to collect and analyze these data. For example, substance abuse affects participants' ability to work; research also suggests that intensive treatment and case management services can help participants with substance abuse issues obtain employment. State employment agencies could collect detailed data on these services and merge it with short-term performance measures to help programs better assess and manage their services. By incorporating this information into a dashboard, program managers and case workers could use it to guide their service delivery strategies, leading to better services and outcomes.

Building a first-class system requires both federal and state participation.

By combining robust administrative data with research techniques such as rapid-cycle evaluations, these systems can help programs test large or small changes, such as new service delivery techniques or phone call reminders, assessing effectiveness in a matter of months rather than years. Predictive analytics can use administrative data to identify patterns and outcomes that suggest approaches to improved program management. When embedded in case management systems, predictive analytics can even help caseworkers identify the best services for an applicant or recognize when information on an application, such as income, may be suspect.

LOOKING AHEAD

Successfully implementing a performance management system requires careful thought to create an accurate, comprehensive, transparent, standardized, stable, and useful structure. Lessons learned from Mathematica's performance management experience suggest that building a first-class system requires both federal and state participation.

At the federal level, agencies must carefully develop measures that fully incorporate and assess key aspects of the program. They must review all measures to identify gaps and create specifications stating how to calculate the measures, including the associated data element specifications. With this information in hand, agencies can develop a web-based application to calculate state performance measures and assess data quality. Lastly, training and technical assistance can ensure that states have the know-how to collect the required data, use the reporting system, and understand the results.

At the same time, states must be engaged in the process. State involvement ensures that federal agencies understand the consequences of their decisions and implement an accountability system that states are prepared to use. Just as importantly, states must recognize that federal systems focus on federal accountability, not on the information that states need to manage and improve their programs. States must consider their internal management and reporting needs as they plan and implement performance management systems. Thinking creatively about the data and measures needed to manage and evaluate their programs, and exploring new research and quantitative techniques that can help them put data to better use, can reveal promising approaches to case management and service delivery.