

Draft, October 2004

Integrated Performance Information for Workforce Development

A Blueprint for States

Draft: October 2004

Washington State Workforce Training and Education Coordinating Board
On behalf of the
Integrated Performance Information Project State Teams from Florida, Michigan, Montana,
Oregon, Texas, and Washington
Principal author: Bryan Wilson, Ph.D.
bwilson@wtb.wa.gov

Integrated Performance Information for Workforce Development: A Blueprint for States **Table of Contents**

Executive Summary	i
Introduction	1
Section One: Challenges and Responses	5
Introduction	5
Authority	5
Culture of Shared Accountability	7
Trust	10
Capacity	11
Funding	11
Privacy	12
Access to Unemployment Insurance Wage Records	14
Reaching Consensus on Goals and Measures	15
Section Two: IPI Performance Measures	17
Introduction	17
What Do Policy Makers Want to Know About Performance?	18
What Makes for a Good Performance Measure?	20
IPI Performance Measures	22
Results for Employers and the Economy	23
Labor Market Results for Program Participants	26
Skill Gains	27
Return on Investment	29
Data Sources and Supplemental Data	31
Performance Targets and Consequences	32
Adjustments	34
Section Three: Shared Information Systems	36
Introduction	36
Authorization and Governance	37
Purpose of the Data Warehouse	38
Leadership and Oversight	38
Funding	39
Management and Operation	40
Participating Agencies/Programs	41
Data Sharing Agreements	41
“Ownership” of the Merged Database	41
Data Access	42
Confidentiality	42
Information Reporting	43
Editorial Review Policies	44
System Modifications	45
Managing Information Flows	45
Data Documentation	46
Quality Assurance	47
Data Transmittal, Storage and Archiving	48
The Sequence for Addressing Issues	49
Conclusion	50

Integrated Performance Information for Workforce Development A Blueprint for States

Executive Summary

Introduction

This is a guide for states interested in creating or further developing integrated performance information for workforce development programs. Integrated performance information reports performance results consistently across programs, across levels (from institutions to local areas to states), or for programs as a system.

A shared information system that produces reports based upon consistent protocols and measures is one key element of integrated performance information. Integrated performance information, however, is more than an information system and a set of consistent measures. It also requires institutions and practices that support shared accountability for results. The Blueprint discusses each of the steps involved: establishing authority, building a culture of shared accountability and trust, generating capacity, crafting performance measures, setting and using targets, as well as creating and maintaining a shared information system.

There are many advantages to states having integrated performance information. The advantages include increased accountability, improved strategic planning, better research, efficiency, and a sense of shared-responsibility among workforce development programs. These advantages can improve the credibility of workforce programs and, in turn, enhance the support they receive and their ability to serve customers.

What is meant by workforce development? Workforce development programs are programs that prepare people for employment and career advancement throughout their lives, and include, but are not limited to:

- Secondary Career and Technical Education
- Postsecondary Career and Technical Education
- The Employment Service, Workforce Investment Act (WIA) Title III
- WIA Title I-B
- Adult Education and Family Literacy, WIA Title II
- Vocational Rehabilitation, WIA Title IV
- Temporary Assistance for Needy Families Work Program
- Apprenticeship

Viewed as a system, most money for these programs comes from the states. The largest programs, in terms of funding, are the education programs for which states typically supply at least 90 percent of the funds. Given this funding arrangement, it is logical that states exercise leadership in devising integrated performance information across workforce development programs.

The United States Office of Management and the Budget (OMB) has taken an important initial step toward integrated performance information by issuing “common measures” for federal workforce development programs. OMB has gotten the attention of federal agencies, particularly the Department of Labor (DOL), regarding the need for consistent measures across

programs. DOL has also taken the step of designing a new reporting system, ETA Management Information and Longitudinal Evaluation System (EMILE), that is to be consistent across most Department programs. This Blueprint builds on these initial steps by recommending performance measures and an information system that would support consistency across state as well as federal workforce programs.

The Blueprint was produced through the joint efforts of six states (Florida, Michigan, Montana, Oregon, Texas, and Washington), with the financial support of DOL. Washington State's Workforce Training and Education Coordinating Board (Workforce Board), with the assistance of the NGA's Center for Best Practices convened policy and technical teams from each of the six states. Each state team included representation from a cross-section of workforce development programs. This diversity of representation was a necessary ingredient for the success of the project. The team members endeavored to listen closely to the perspectives of each program and to arrive at solutions that were acceptable to all.

The state teams met several times during 2004 to share experiences and lessons learned, review technical papers, think through key questions, and arrive at a consensus on key aspects of integrated performance information. In addition to the NGA's Center for Best Practices, the states received assistance from the Ray Marshall Center at the University of Texas and The Center for Governmental Studies at Northern Illinois University. The states benefited greatly from the research conducted on behalf of the project as well as from the general expertise and experience of these entities. They and the state teams provided much of the material for the Blueprint and reviewed and commented on drafts. The Blueprint is very much the shared product of the six states and their partners, although the Washington Workforce Board remains ultimately responsible.

The Blueprint consists of the following sections:

Challenges and Responses

States face serious challenges as they embark down the road of creating integrated performance information. And the challenges don't end with the beginning. Most of the challenges are ongoing and require constant attention. This section of the Blueprint examines some of the major challenges and some choices that states have made to address them. The section discusses: (1) establishing authority for integrated performance information, (2) creating a culture of shared accountability, (3) building capacity—including funding and addressing privacy issues, and (4) reaching consensus on goals and measures. For each challenge, the Blueprint presents some examples of actions states have taken, a variety of actions reflecting the variety of institutional structures and political conditions in states. States reading the Blueprint may want to pick and choose from these examples and implement the steps that best fit their situation.

IPI Performance Measures

What are the best performance measures for workforce development if the same measures are applied horizontally and vertically within the system, including programs that are funded mostly by the states and programs that are funded mostly by Congress? This was a central question considered by the six states and their partners at NGA's Center for Best Practices and the Ray Marshall Center at the University of Texas. The states began by considering what policy leaders

want to know about performance results. They then suggested the selection criteria for judging measures, and analyzed the advantages and disadvantages of a long list of possible measures. In the end, the states agreed on a relatively short list of measures that best respond to the performance questions commonly posed by policy leaders. This section of the Blueprint summarizes the discussion and recommendations of the six states' teams. Other states may wish to follow suit, either by following this type of process within their state, or by adopting some or all of the recommended measures. The following table summarizes the performance measures recommended by the teams from the six states. The measures are separated into those measures that are useful as accountability measures, for which there could be targets and consequences, and those measures that, while indicators of how well the workforce development system is doing, do not sufficiently satisfy the criteria for good performance measures in order to be used for targets and consequences.

IPI Performance Measures

Category	Measure
Accountability Measures	
Labor Market Results for Program Participants	<p>1. Short-term Employment Rate: The percentage of participants who have exited with employment during the second quarter after exit. (For youth, enrollment in education counts as well as employment.)</p> <p>2. Long-term Employment Rate: The percentage of participants that have exited with employment during the fourth quarter after exit. (For youth, enrollment in education counts as well as employment.)</p> <p>3. Earnings Level: Median earnings during the 2nd quarter after exit among all exiters with any earnings.</p>
Skill Gains	<p>4. Credential completion rate: The percentage of exiters who have completed a certificate, degree, diploma, licensure, or other industry-recognized credential within one year of exit.</p>
Results for Employers and the Economy	<p>5. Repeat Employer Customers: The percentage of employers who are served, directly or indirectly, who return for the same service.</p>

Performance Indicators	
Results for Employers and the Economy	6. Employer Market Penetration: The percentage of all employers who benefit from the services.
Return on Investment	<p>7. Taxpayer Return on Investment: The present value of the net impact on tax revenue and social welfare payments compared to the cost of the services.</p> <p>8. Participant Return on Investment: The present value of the net impact on participant earnings compared to the cost of the services.</p>

The state teams considered many other measures beside these eight. This section discusses some of these measures and why they did not rise to the top of the list. This section also discusses data sources, setting and using performance targets, adjusting targets or results for economic conditions and participant characteristics, and how national research complements state performance measurement.

Shared Information Systems

A major barrier to integrated performance information is the absence of a shared or integrated information system for workforce development in most states. Participant information remains scattered about in various program level Management Information Systems (MISs), and there is no system for integrating the information from the multiple MISs to support performance measurement and reporting.

This section of the Blueprint discusses the steps and decisions involved in establishing a data warehouse that links administrative records from multiple programs with other data sets containing outcome information, such as unemployment insurance wage records. A data warehouse is built on top of exiting MISs and does not replace them. The MISs are still necessary for program management. The warehouse is a longer-term repository where data are cleaned and matched in order to analyze and report performance outcomes.

Creating a data warehouse requires states to make decisions regarding: authorization; leadership; funding; scope; data ownership, confidentiality, and access; information flows; reporting; and quality assurance, among other issues. This section of the Blueprint walks through each of these issues and the major options.

Conclusion

States will want to consider how to use this Blueprint given federal initiatives in this area, including the OMB's common measures, the DOL's EMILE System, and the pending reauthorization of the Workforce Investment Act (WIA), the Carl D. Perkins Vocational and Technical Education Act, and other federal acts related to workforce development. States may choose to proceed in a number of different ways. The development of a shared information system is something that should be useful no matter the outcomes of the federal initiatives. Electronically linking records from multiple programs with files containing outcome data will facilitate the implementation of common measures and reporting the performance information likely to be required by the reauthorized acts, as well as meeting state-identified needs.

If states find some or all of the performance measures recommended here to be useful, states can implement them as additional measures to those necessary to satisfy federal requirements. Experience has shown, moreover, that federal performance measurement requirements evolve over time. If states find the IPI measures to be useful, the measures may be reflected in future generations of federal acts and guidelines. Federal performance requirements have a powerful effect on program implementation and results; if states find the IPI measures to be useful, it would be very helpful if future federal requirements were aligned with the measures.

Finally, no matter the specific course of events in Congress or the federal agencies, the basic issues of building the capacity for and a culture of shared accountability are likely to be challenges that remain with states. Hopefully, the state experiences and lessons learned that are shared here will help states as they continue to face these challenges.

Integrated Performance Information for Workforce Development

A Blueprint for States

Introduction

This is a guide for states interested in creating or further developing integrated performance information for workforce development programs. Integrated performance information reports performance results consistently across programs, across levels (from institutions to local areas to states), or for programs as a system. Examples include: using the same measure to report both employment rates for a radiology program at a community college and employment rates for the state's two-year colleges as a whole; One-Stop center measures whereby the results for One-Stop participants are summed up without regard to program silo; and measures that indicate the results for taxpayers or employers that are produced jointly by multiple programs.

A shared information system that produces reports based upon consistent protocols and measures is one key element of integrated performance information. Integrated performance information, however, is more than an information system and a set of consistent measures. It also requires institutions and practices that support shared accountability for results. The Blueprint discusses each of the steps involved: establishing authority, building a culture of shared accountability and trust, generating capacity, crafting performance measures, setting and using targets, as well as creating and maintaining a shared information system.

There are many advantages to states having integrated performance information. The advantages include increased accountability, improved strategic planning, better research, efficiency, and a sense of shared responsibility among workforce development programs. These advantages can improve the credibility of workforce programs and, in turn, enhance the support they receive and their ability to serve customers.

Accountability for performance is a banner raised by virtually every elected official. Elected officials want programs to report their results in understandable language so the officials can judge whether or not the programs are good investments. Elected officials are also very busy people. They do not have time for someone to explain dozens of different measures, each with a different denominator, and multiple definitions for seemingly straightforward terms such as employment and earnings. A core of relatively simple performance measures with consistent definitions and methodology is a great tool for convincing elected officials that programs track how they are performing and are accountable for results. Without such evidence, programs cannot expect legislative or budgetary support, particularly during tight fiscal times.

Workforce development programs complement one another in serving many of the same customers. The same individual may attend secondary and postsecondary career and technical education, use labor market information at a One-Stop, and find a job through the Employment Service. The results are the shared effect of the services of all these programs. The results will be better and most efficient when states strategically plan the services of multiple programs in an integrated fashion. Having consistent performance measures across programs, and shared information to support them, greatly facilitates such planning.

Providing good service to customers depends upon knowing what works and what needs to be improved. Integrated performance information enables analysis of the relationship between services and results. For example, a comprehensive system can perform a longitudinal analysis of what happens to customers over time, regardless of funding source, and provide breakdowns by demographic characteristics and geographic location. One can detect the patterns of service and program participation that work best for different sets of customers. Such an analysis is not very practical without consistent performance measures and a shared information system.

Integrated performance information is also more efficient—by integrating aspects of separate reporting systems redundant efforts can be eliminated. Moreover, the electronic linking of program records with outcome data can replace more expensive methods such as surveys.

The advantages of integrated performance information for public programs are perhaps nowhere greater than in the area of workforce development. For many years, observers have widely commented on the system's complexity; how there are many programs with insufficient coordination among them causing inconvenience to customers and confusion to planners and policy makers. Integrated performance information can improve coordination, enabling programs to act more as a system even though they may remain separately administered by agencies.

There is also a reputation in workforce development, some deserved, some undeserved, of questionable results. Reports of weak performance for certain programs have, in the minds of some, seeped over to other programs unfairly tainting the whole enterprise. (This is another example of how the programs share responsibility for results.) Integrated performance information is a tool that can demonstrate the performance of the system, showing where performance is strong, and informing improvements where performance is relatively weak.

What is meant by workforce development? Workforce development programs are programs that prepare people for employment and career advancement throughout their lives, and include, but are not limited to:

- Secondary Career and Technical Education
- Postsecondary Career and Technical Education
- The Employment Service, Workforce Investment Act (WIA) Title III
- WIA Title I-B
- Adult Education and Family Literacy, WIA Title II
- Vocational Rehabilitation, WIA Title IV
- Temporary Assistance for Needy Families Work Program
- Apprenticeship

Viewed as a system, most money for these programs comes from the states. The largest programs, in terms of funding, are the education programs for which states typically supply at least 90 percent of the funds. Given this funding arrangement, it is logical that states exercise leadership in devising integrated performance information across workforce development programs. Even for the programs that are funded mostly by Congress, such as Workforce Investment Act (WIA) Title I-B, the actual skills training is provided mostly by state supported community and technical colleges. The federal funds usually pay for the cost of tuition, fees, and books, while the states pick up the larger share of the colleges' cost through their general funds.

The United States Office of Management and the Budget (OMB) has taken an important initial step toward integrated performance information by issuing “common measures” for federal workforce development programs. OMB has gotten the attention of federal agencies, particularly the Department of Labor (DOL), regarding the need for consistent measures across programs. DOL has also taken the step of designing a new reporting system, ETA Management Information and Longitudinal Evaluation System (EMILE), that is to be consistent across most Department programs. This Blueprint builds on these initial steps by recommending performance measures and an information system that would support consistency across state as well as federal workforce programs.

Most states do not have integrated performance information for workforce development programs. There is not consistent information across programs, across levels, or for workforce development as a system. Governors and legislators, their staff, and other policy leaders cannot easily make heads or tails of all the different federal and state performance measures, or how the programs use information on results to improve performance. Most states do not have shared information systems that can produce consistent information on results, provide longitudinal data, or slice and dice data in different ways as needed for policy analysis.

Some states are at just the initial stage of considering whether they want integrated performance information, others may have been at it for a long time, but are interested in improving their efforts based upon the experience of other states. In either case, this Blueprint is intended to be of assistance. States may want to consider bits and pieces, or the whole thing, as best suits their needs.

The Blueprint was produced through the joint efforts of six states (Florida, Michigan, Montana, Oregon, Texas, and Washington), with the financial support of DOL. Washington State’s Workforce Training and Education Coordinating Board (Workforce Board), with the assistance of the National Governors Association’s (NGA) Center for Best Practices convened policy and technical teams from each of the six states. Each state team included representation from a cross-section of workforce development programs. This diversity of representation was a necessary ingredient for the success of the project. The team members endeavored to listen closely to the perspectives of each program and to arrive at solutions that were acceptable to all.

The state teams met several times during 2004 to share experiences and lessons learned, review technical papers, think through key questions, and arrive at a consensus on key aspects of integrated performance information. In addition to NGA’s Center for Best Practices, the states received assistance from the Ray Marshall Center at the University of Texas and The Center for Governmental Studies at Northern Illinois University. The states benefited greatly from the research conducted on behalf of the project as well as from the general expertise and experience of these entities. They and the state teams provided much of the material for the Blueprint and reviewed and commented on drafts. The Blueprint is very much the shared product of the six states and their partners, although the Workforce Board remains ultimately responsible.

The Blueprint's framework closely follows the discussions of the six states. It covers the challenges states face in creating integrated performance information, and how leading edge states have responded to these challenges. It talks about unclear authority, the lack of a culture of accountability, insufficient capacity, and a lack of consensus on goals and measures. The Blueprint offers state experiences in handling these challenges as options that states may want to consider, rather than as the right answers or the only paths to success.

The Blueprint discusses what policy leaders want to know about performance results, and selection criteria for choosing the best measures to answer their questions. The Blueprint presents the performance measures that these six states came to agree are the best measures for integrated performance information. It also covers issues of setting performance targets, using regression models, establishing consequences for results, and how state performance measurement and national research complement one another.

Finally, the Blueprint talks about shared information systems. The Blueprint presents the basic issues that states must consider when building shared information systems, including governance, funding, data ownership, reporting, and data dissemination. The Blueprint assumes that states will produce integrated performance information by linking electronic administrative records maintained in multiple management information systems (MISs), as opposed to building a huge single MIS. The prime example of such linking is matching participant records with unemployment insurance wage records in order to create data on participant employment and earnings. An information system that performs such linking for multiple programs is what is meant by a shared information system.

One last note before beginning, the Blueprint focuses on policy level information—the type of information needed by elected officials and state directors of programs and their policy staff. It is not designed for those seeking day-to-day management information. As discussed later in this document, pp. ____, these are two different audiences with very different sets of needs that require different solutions. While this gap in the Blueprint may be disappointing to hardworking staff at the ground level, without the support of policy leaders, their good work would not be possible.

Challenges and Responses

Introduction

States face serious challenges as they embark down the road of creating integrated performance information. And the challenges don't end with the beginning. Most of the challenges are ongoing and require constant attention. This section of the Blueprint examines some of the major challenges and some choices that states have made to address them. The section discusses: (1) establishing authority for integrated performance information, (2) creating a culture of shared accountability, (3) building capacity—including funding and addressing privacy issues, and (4) reaching consensus on goals and measures. For each challenge, the Blueprint presents some examples of actions states have taken, a variety of actions reflecting the variety of institutional structures and political conditions in states. States reading the Blueprint may want to pick and choose from these examples and implement the steps that best fit their situation.

I. Authority

Given the fragmentation of workforce development programs, there is generally no clear authority for any one entity to lead integrated performance information for the workforce system. The multiplicity of silos among federal programs is well documented. Three different federal agencies (Labor, Education, and Health and Human Services) are responsible for administering the largest programs, and there is no standing interagency body to coordinate their efforts. The fragmentation begins in Congress where different committees deal with labor and education programs on the one hand, and social welfare programs on the other.

At least in Congress, the same committee in the House and Senate deal with both education and labor programs. This is usually not the case in states; where frequently there are separate legislative committees for K-12 education, higher education, labor, human services, and economic development, each with a piece of the workforce development puzzle. Programs tend to bear the stamp of one committee or another, and policy areas and issues that span committee jurisdictions are not always well managed.

Governors have often taken the lead in addressing fragmentation among workforce development programs. They have followed one of three paths: (1) consolidating workforce development programs into a single agency, (2) establishing a coordinating body, or (3) attempting coordination through interagency agreements. Whichever path is followed, there needs to be a champion of integrated performance information in order to make it happen, and most states have found that the clearer and stronger the authority that is established, the more likely it is to last and for the state to succeed.

In Washington, the Governor requested legislation to create the Workforce Training and Education Coordinating Board (Workforce Board). Established in 1991, the Workforce Board is an independent agency that does not operate any of the major programs. Instead, it focuses on policy planning and performance accountability. One of the primary motivations behind its creation was to establish a consistent performance accountability system for workforce development programs. The intent section of the Workforce Board's authorizing statutes state, "The work force training and education system's data and evaluation methods are inconsistent and unable to provide adequate information for determining how well the system is performing

on a regular basis so that the system may be held accountable for the outcomes it produces.” The statutes give the Workforce Board the authority to establish common definitions, common standards for data collection and maintenance, common standards for program evaluations, and the assignment to conduct evaluations, including net impact and cost benefit evaluations, of the workforce development system. This authority was of keen interest to the business and labor communities who strongly felt, and correctly so, that the state did not have a good handle on how well its workforce programs were performing. They championed these provisions before the legislature.

The legislation also gained the support of the administrative agencies, in part, because the Workforce Board would not itself administer programs and the largest administrative agencies (and only the largest) would each have one vote on the Board (with business and labor having a voting majority). The administrative agencies believed this structure would enable the Workforce Board to neutrally establish standards that would create a level playing field among the major programs. They also believed that independence from program operation would give the Workforce Board credibility when reporting performance results to the Governor and Legislature. (See Appendix ___ for a copy of the Workforce Board’s statutes.)

In Texas, two paths have been followed. Texas has both merged agencies and created a coordinating body with clear responsibility for establishing integrated performance information. In 1993, Texas established in statute the Texas Council on Workforce and Economic Competitiveness with responsibility for strategic planning and oversight for the state’s workforce development programs. In 1995, again by legislation, Texas merged two dozen workforce programs into a single new agency, the Texas Workforce Commission. The legislation also maintained the Council on Workforce and Economic Competitiveness, moved the Council into the Governor’s Office, and charged the Council with establishing comprehensive, system-wide performance measures.

In 2001, Texas enacted legislation providing additional guidance to the Council, now named the Texas Workforce Investment Council, regarding the development of performance measures. This legislation suggests in general terms some possible measures, including job placement rates, job retention rates, and wage rates. (See Appendix ___ for a copy of the Texas statute.) The legislation reinforces the authority of the Workforce Investment Council by clearly stating, “The council shall include in the strategic plan goals, objectives, and performance measures for the workforce development system that involve programs of all state agencies that administer workforce programs.” The Governor has supported the efforts of the Council, and approved the performance measures developed by the Council, with the Governor’s Office of Budget, Planning and Policy playing a key role.

Oregon has established a Governor’s Workforce Policy Cabinet. The Cabinet includes the agency directors for workforce development, economic development, and education, in addition to the Governor’s own workforce and education staff. In a sense, the cabinet is a virtual alternative to consolidating agencies. The cabinet discusses policy issues, including systemwide indicators, to form a consensus and agree on action.

Florida has followed another path. The Governor and Legislature created Florida’s integrated performance information system prior to establishing a broader coordinating body for workforce development. In 1982 the Governor’s staff prompted an initial interest in student follow-up analysis using unemployment insurance wage records through its 1982 study of vocational

education. The study concluded that the available information was inadequate, especially information supporting analysis of labor market outcomes. As a result, the legislature created the Occupational Identifier Project in 1984. The Occupational Identifier Project used unemployment insurance wage records to measure post-program employment. At the same time, the legislature enacted a performance requirement for secondary and postsecondary vocational education that 70 percent of any program's completers had to be positively placed as a condition of continued funding. Initially, placement could be collected through local mail or telephone surveys or through the Occupational Identifier Project. In 1988 as a result of the Project's progress, the legislature enacted Section 229.8075 of Florida Statutes creating the Florida Education and Training Placement Information Program (FETPIP) as a special unit within Florida's education agency. FETPIP then became the primary resource for documenting compliance with the vocational education performance requirement.

Since 1988, FETPIP has expanded its follow-up data coverage and its services to education and training organizations, including program evaluation, performance-based funding, and consumer information. FETPIP follow-up data includes employment and earnings, postsecondary education, military enlistment, incarceration, and welfare participation. FETPIP produces information for:

- ❖ Reviews of programs of study at four-year colleges and universities.
- ❖ Florida's performance-based budgeting system.
- ❖ Workforce Investment Act performance indicators.
- ❖ Carl D. Perkins performance indicators.
- ❖ Adult Education and Family Literacy performance indicators.
- ❖ K-12 school reports.
- ❖ The Legislative Office of Program Policy Analysis and Government Accountability.
- ❖ Florida's K-20 Education Performance Accountability System.
- ❖ Consumer tools such as college program catalogues.
- ❖ Research, including return on investment studies, the costs of dropping out of high school, and studies of the education pipeline.

The scope of the FETPIP statute has been broadened over the years to reflect these additional functions.

II. A Culture of Shared Accountability

Legal authority by itself does not necessarily make things happen. Statutes are full of provisions that do not mean a thing, either because of resistance or a lack of interest. Building a culture of accountability—patterns of behavior and beliefs that are truly supportive of accountability—takes a great deal of time and effort. This is especially true when creating a culture of *shared* accountability that crosses many programs. Workforce development programs have long been held accountable for their own performance, dating from the introduction of the first federal performance standards over two decades ago. Getting programs to pay serious attention to the performance of other programs and of the system as a whole, and to recognize the benefits of an integrated system is another matter.

Performance measurement requirements that span multiple programs involve giving a third party authority to set standards for others. A program may be a part of that entity if it is board or council but so will others, and their views are not going to be identical. There is inherently

resistance to another entity's authority to establish performance accountability requirements that include one's own program. There will be fear that another entity will misuse its authority to one's disadvantage. This may be due not to any deliberate intent to harm, but simply due to insufficient understanding of the purpose of the program and the customers it serves. In order to succeed there must be a building of trust over time that authority will not be misused. And in order to obtain active participation, programs need to see something in it for themselves—a benefit from participating—rather than viewing it as just another compliance activity.

States have employed a number of strategies to create a sense of shared accountability for performance results. These efforts have included both large-scale state initiatives and efforts based on more narrowly appealing to programs' self-interest.

In the case of Oregon, the effort began with a statewide strategic planning effort championed by the Governor—"Oregon Shines." Oregon launched this planning process in 1988 in response to a severe recession. The focus was, and continues to be on economic vitality, and workforce development is a major component. The first of the three key goals of the original plan was: "A superior work force: Invest in Oregonians to build a work force that is measurably the most competent in America by the year 2000, and equal to any in the world by 2010." The process began with 16 committees composed of 180 business, labor, education, and government leaders. Input was obtained from literally thousands of people throughout the state. Among the committees' recommendations was the creation of the Oregon Progress Board, chaired by the Governor, to carry on their work. The Progress Board was created with the responsibility "to remind us of our shared vision, monitor our progress in achieving measurable goals, and bring choices to our attention."

As part of this strategic planning process, Oregon developed the "Oregon Benchmarks," initially 259, now 100, measures of success. The benchmarks include goals that are not the responsibility of any one program or agency, and results that go beyond outcomes for program participants. According to the Oregon Progress Board, the benchmarks have been "an excellent tool for encouraging collaboration among different interests, engendering long term thinking and developing results oriented management systems."

In 1997, the Progress Board issued "Oregon Shines II." The benchmarks related to workforce development include:

- ❖ The percentage of high school students that have completed a structured work experience: The goal for 2010 is 100 percent.
- ❖ The percentage of Oregon adults (age 25 and older) who have completed an associate degree in professional-technical education: The goal for 2010 is 10 percent.
- ❖ The percentage of Oregonians in the labor force who received at least 20 hours of skills training in the past year: The goal for 2010 is 100 percent.

The benchmarks are the broadest level of three tiers of performance measures. The most specific of the three tiers is program performance measures, including federally required program measures and other measures that are reported to the state legislature. Since 1993, the legislature has required agencies to establish program measures linked to the benchmarks, and to include the measures as part of the agency biennial budget submittals. The middle of the three tiers are system-wide measures for workforce development, looking at results of the primary One-Stop partners (see p. ____).

This three-tiered system of performance measures has supported a pervasive culture of performance accountability in Oregon. It has helped to create a focus on ultimate results and a sense that “we are all in this together.”

There are other paths to creating a culture of accountability that do not depend upon such a big, statewide effort. Florida’s FETPIP system has created a culture of accountability built upon the individual interest of programs and organizations (not that the state level initiative has been unimportant). FETPIP can calculate many measures providing information of unique value to particular programs. The direct value of the information to particular customers was especially crucial in the beginning stages. After a few years of initial development, FETPIP was approved to proceed with a pilot demonstration project. As a part of the pilot, FETPIP staff worked with three large (and influential) school districts, with postsecondary programs, and one large community college. Staff worked directly with their data and program managers to understand the particulars regarding their data as well as the reasons for collecting the data. As initial matches were done using student records and unemployment insurance wage reports, staff carefully reviewed results with, and incorporated suggestions from, the pilot agencies. As FETPIP staff and local users became more comfortable working with the data, they generated detailed reports that carved the data in as many ways as they could to pique interest and explored new ways of using data to improve programs.

The payoff was that when FETPIP staff needed support with legislative committees, they did not need to lobby for that support on their own. They had created “champions” who could speak well of the effort. Since then, FETPIP has continued this strategy. It is constantly on the lookout for opportunities to apply its services, and thereby create more champions.

Another means of engendering support for shared accountability is more direct financial incentives. While the provision remains controversial among states, the Washington Workforce Board finds the inter-program incentive fund authorized by Section 503 of WIA useful for reinforcing shared accountability. That section authorizes the allocation of incentive funds to states that meet performance targets for each of WIA Title I, Adult Education and Family Literacy, and Carl D, Perkins Vocational and Technical Education. If a state fails to meet the targets for any one of the three, it is ineligible for an incentive award. When Washington receives a section 503 award, the Workforce Board allocates the funds to local areas where all three programs achieve the targets, and local areas may expend the funds only on activities that benefit all three programs. This has helped local programs appreciate the importance of their colleague’s results and enhanced the feeling of shared accountability.

Trust

A culture of shared accountability is built upon trust—trust that performance information will not be used against one's interests. Trust is built step by step.

Washington's Workforce Board began by simply reporting performance. The Workforce Board's strategy was to first get programs used to having another entity report their performance. Other agencies experienced the advantages of having a third party report their performance to the Governor and Legislature—an agency that could have greater credibility since it was not directly responsible for program performance. Targets were established only after four years of performance reporting, and only after the targets were in place for a year were there any financial consequences for performance (through the allocation of relatively small pots of incentive funds).

The Workforce Board staff discusses draft evaluation reports with the relevant agencies before they are published. This gives the agencies the opportunity to correct any mistakes or misinterpretations of the data and to suggest changes in the tone or substance of the narrative. The reports avoid any language that smacks of "gotcha." When numbers are low, programs are not labeled as "bad" (as some legislators have desired). Instead, the reports speak of areas for improvement. And care is taken that the decision criteria used for identifying areas needing improvement are the same across programs.

Regression modeling, another feature of the Workforce Board's accountability system, can be a useful tool for building trust. Programs are usually concerned that their results will look bad if they serve hard to serve customers or if the economy goes down hill. Programs are worried that they will be "punished" for factors outside their control. Regression models and other techniques that adjust either performance targets or results for changes in participant demographic characteristics or economic conditions, can help alleviate these fears.

Florida emphasizes that an important precondition of trust is confidence that the data is good. When data are collected by linking several databases from different agencies, one is dependent on controls that may or may not be applied by others. For example, the accuracy of employment data requires that employers accurately record and report employee identification and payroll information. It also requires that the unemployment insurance agency enters the employers' data accurately and assigns the North American Industrial Classification Codes correctly. Similarly, the accuracy of participant data requires that social security numbers be collected, recorded, verified, and transmitted accurately by frontline staff. Other participant information such as demographics, socio-economic characteristics, and program characteristics, must also be faithfully represented.

For the sake of quality control, those producing integrated performance information by linking records from different sets of administrative data should have a clear understanding of the data sets involved. They should know how the data are originally collected and recorded, how they are processed, how they are defined, and what they represent. As the record linking activity proceeds, any anomalies or problems should be brought to the attention of the source agency for clarification or resolution. Use is often the best means of maintaining data quality. By actually using data and alerting appropriate staff when anomalies are found, data are kept clean and their quality improved over time.

While formal agreements in Florida stipulate many of the circumstances regarding data handling and processing, there is by now a sense of trust that goes beyond formal agreements. Maintaining this trust involves working with agencies to assist them in recognizing and resolving problems that are uncovered in the data matching process. It also means continually working with them to ease any processing burdens associated with the matching effort. Where there are questions regarding the release of certain data or analysis, even if the release appears to fall within the agreement's parameters, the cooperating agency is consulted.

III. Capacity

In addition to authority and will, designing and operating a system of integrated performance information requires the capacity to do so. States must have the financial resources and staff with the required skills to make it happen. They must also have appropriate access to the necessary data in order to consistently measure performance across programs and levels.

Funding

States have expressed concern that integrated performance information will require large investments they cannot make, but states' experiences do not support this concern.

In Texas, workforce development agencies share in the costs of integrated performance information as set forth in state statute enacted in 2001. The statute directs the Texas Workforce Investment Council to establish, with the approval of the Governor, a formula "to determine the level of support each agency administering a workforce program must provide to operate the automated follow-up and evaluation system." The formula developed by the Council is based on the number of seed records evaluated for each workforce training or education program.

In Florida, FETPIP is supported by federal and state general funds, as well as fees for service. FETPIP had rather modest beginnings. In 1984, the legislature appropriated \$60,000 for the initial study and design, enough to support one staff person and some consulting. Over the course of the next several years, support grew, including funds from Wagner-Peyser and the Job Training Partnership Act, supporting a staff of four. In 1988, the legislature created FETPIP in statute and provided for its staffing, including a director, two senior data base analysts, two research assistants, and clerical support. Its annual budget was about \$350,000. Beginning in 1994, FETPIP began collecting some data on a quarterly basis (before data collection was annual). FETPIP's budget and staff increased accordingly.

At the current time, FETPIP has nine professional level permanent positions and one full-time clerical staff. The professional positions include a director, two senior data analysts, two systems analysts, three junior level computer programmers and one research assistant. The annual budget includes approximately \$400,000 in state general revenue resources, \$200,000 from the Agency for Workforce Innovation (from a combination of federal WIA, Wagner-Peyser, and TANF resources), and \$60,000 from the Perkins Act. FETPIP also receives funding through fees for services.

The Workforce Board relies on three funding sources to support its integrated performance information system: state general funds, the Governor's 10 percent funds for WIA Title I state-wide activities, and Carl D. Perkins Administrative funds. Serving as Washington's sole state agency for Perkins as well as the state workforce investment board for WIA, places the

Workforce Board in a unique position to access these three funding sources, providing a total of approximately one million dollars per year to support six FTEs and contracted services. These resources provide a variety of services in addition to data matching and reporting, including maintaining the state eligible training provider list, extensive surveys of participants and employers, net impact and cost benefit evaluations, and periodic research projects for a state workforce development system expending close to 900 million dollars per year.

Privacy

Integrated performance information relies upon access to administrative records in order to have consistent and affordable data. A paramount issue in gaining access to administrative records is the protection of individual privacy. In order to have this capacity, certain precautions are essential. Without taking appropriate precautions to protect privacy, a state will quickly find that it has no capacity to produce integrated performance information.

The issue of privacy is heightened because the only data element that uniquely identifies unit records that is common to most data sets, including the very important Unemployment Insurance (UI) wage records, is the social security number. With identity theft a growing problem, individuals are understandably concerned about sharing their social security number. States have employed several methods in order to protect individuals from the state or its agents improperly divulging their number (see p. ____).

Laws and regulations restrict the conditions under which government agencies may share individual information with others. Many programs have their own unique restrictions that must be respected when establishing integrated performance information. Some of the most important restrictions, although by no means the only ones, are those associated with student and unemployment insurance records.

In the use of student records, states must satisfy the requirements of the Family Educational Rights and Privacy Act (FERPA). FERPA protects the privacy interests of students and/or the parents of students who are minors with respect to their personal education records. FERPA allows students, or their parents/guardians, if they are minors, the right to review and inspect their educational records, the right to amend them, and the right to have some control over their disclosure. In general, personal educational records cannot be disclosed without the prior written consent of the student (or his or her parent or guardian if a minor). Exceptions are provided if:

- The disclosure involves directory information. This is information in a personal education record which would not be considered harmful if it were released. Directory information includes such items as student name, address, telephone number, date and place of birth, major field of study, degrees or certificates awarded, and others that may be specified in state law.
- The disclosure is to school officials, including teachers, who have been determined to have legitimate educational interests.
- The disclosure is to organizations conducting studies on behalf of educational agencies or institutions to validate tests, administer student aid programs, or improve instruction.

The use of student records to create performance information for program reviews and accountability comes under the exception of disclosure to improve instruction. The organization receiving the student records must conduct the process in a manner that precludes the use of personally identifiable information about students or their parents by individuals other than employees that represent the receiving organization. The receiving organization must also destroy the personal education records when the purposes for which they were provided are completed.

On January 18, 2001, the U.S. Department of Education (DOE) issued a program memorandum concerning the applicability of FERPA to certain provisions in the Carl D. Perkins Act (Perkins) and WIA. The letter recognized that only Social Security numbers can be used to link educational records to UI wage records and outlined options for how this could be accomplished.

On January 30, 2003, DOE issued a new letter outlining three ways that education agencies can exchange student records with state unemployment insurance agencies in order to meet Perkins and WIA performance measurement requirements.

1. The custodial organization of the education records may receive the employment records and link the two sets of records to determine the employment status of former students.
2. Students (or the parents of minors) may grant prior consent for the release of their social security numbers for the purpose of matching their records with employment records for statistical research purposes.
3. The education agency may share the student records with the agency holding the employment records, and that agency may match the two sets of records *if the education agency provides in-person oversight and direction*.

States have used a variety of approaches to meet FERPA requirements. In Florida, data matching is conducted by FETPIP, a unit within the Florida Department of Education. The DOE is the custodian of student records for all public education entities in the state. FETIP negotiated an arrangement with DOL and Employment Security that established a UI wage record archive within FETPIP. FETPIP can thereby match student records with wage records without releasing the student records to an organization outside of education. In exchange for receiving the wage records, FETPIP handles any requests for statistical matches with the UI wage records.

In Texas, when the recent, more restrictive interpretation of FERPA was handed down by DOE, the Texas education agencies (Texas Higher Education Coordinating Board and Texas Education Agency) were no longer able to share their student records with the Texas Workforce Commission that holds the UI wage records. This forced Texas leaders to rethink how they conducted their follow-up and evaluation effort. Currently, the Texas Workforce Commission and the Texas Higher Education Coordination Board jointly administer the follow-up and evaluation system through a Memorandum of Understanding. Under this agreement, the

Commission matches participant records from non-education programs with employment wage records (and federal employment databases) and forwards the results to the Texas Higher Education Coordinating Board, which then matches the records with college master enrollment files. The Texas Higher Education Coordinating Board matches all education records that fall under FERPA with the employment wage records and federal employment databases. The resulting program data is analyzed and returned to Texas Workforce Commission.

Access to Unemployment Insurance (UI) Wage Records

Another data access challenge is access to UI wage records. State unemployment insurance statutes typically limit access to the records provided by employers and job seekers. Sometimes these limits impinge on the state's capacity to use the records to document labor market outcomes for any programs other than unemployment insurance.

In Washington, the Legislature enacted legislation requested by the Employment Security Department to clarify when DOE may share administrative records with other government agencies and the steps required for safeguarding the confidentiality of the records (see Appendix ___ for a copy of the statute). The legislation specifies that DOE may share information provided by employers to other agencies for the purpose of conducting statistical analysis, research, and evaluation studies. More broadly, the statute authorizes DOE to enter into data sharing contracts with other agencies for the purpose of evaluating and improving the operation of state programs. The same statute also specifies the conditions under which job seeker data may be shared with other agencies. DOE may enter into contracts to share job seeker data with other agencies when the data is necessary for the "efficient provisions of workforce programs, including but not limited to public labor exchange, unemployment insurance, worker training and retraining, vocational rehabilitation, vocational education, adult transition from public assistance, and support services" (Revised Code of Washington 50.13.060). The statute provides for a civil penalty of \$5,000 for the misuse or unauthorized release of records or information.

Given the mobility of the labor force, states must also have access to wage records of other states. For purposes authorized under WIA, the Wage Record Interchange System (WRIS) enables states to determine the employment and earnings of their program participants in other states. WRIS, however, is currently limited to WIA. This restriction is a serious impediment to the implementation of integrated performance information—it precludes consistent reporting of performance across programs. A state may, for example, include in consumer reports authorized by WIA out-of-state employment and earnings information for a college training program. The state, however, is not permitted to use this same information when reporting to DOE the performance of the state's postsecondary system. Access to WRIS is also important for consistency across states. There is a great deal of variation among states in the extent of out-of-state employment of former program participants. Without access to WRIS, this variation precludes valid interstate comparisons of employment and earnings results.

Finally, states must also have access to federal employment records held by the Department of Defense and the Federal Office of Personnel Management. Individual states have successfully established agreements with these agencies for the purpose of tracking their former program participants. But it is cumbersome for each of the 50 states to separately approach the two agencies. Currently, DOL's Federal Employment Data Exchange System (FEDES) project is developing a national system for accessing Defense and federal personnel records. It will be important for this effort to include access for programs beyond WIA.

IV. Reaching Consensus on Goals and Measures

Another challenge faced by states is reaching consensus on goals and measures. While goals and measures are sometimes established in statute, probably more often than not the specific measures are left to administrative action. In either event, it is paramount that the stakeholders reach agreement on the common goals they are trying to achieve and the best measures that apply across the workforce development system.

Oregon's systemwide indicators had their beginnings in the Oregon Workforce Option, a partnership among federal, state, and local workforce programs in the early 1990's to reduce regulatory barriers in exchange for increased performance. Oregon Workforce Option partners met with federal representatives (DOL, HHS and DOE) with the aim of replacing the myriad federal measures with a single set of performance measures. At the time, Oregon workforce partners revealed they collectively had 144 individual program measures. A workgroup was formed to craft a single set of measures by "rolling up" the federal and state measures into one set. For example, most of the partners had an outcome measure that was equivalent to "entered employment," even if it was called something else.

The Oregon Workforce Policy Cabinet formed a workgroup called the Performance Accountability Policy group to fully implement the indicators. This group was comprised of both policy staff and technical experts from each of the affected agencies, plus local representatives. The policy group developed 13 performance indicators that were outcome focused, measurable, based on available data, and inclusive of the work of all the partners. The policy group recognized that for the agencies to be involved, they needed to have their work reflected in the indicators.

A parallel track was developed with the Oregon Workforce Investment Board. The co-chairs of the policy group staffed the Workforce Investment Board's Performance Accountability Committee. As soon as the indicators were fully developed, they were discussed and approved by the Workforce Investment Board. This made the state board the "owner" of the systemwide indicators.

In Washington, the Workforce Board's statutes do not specify the goals or measures for the workforce development system; the statutes give the Workforce Board the assignment to establish the goals and measures "in coordination with the operating agencies." To do this, the Workforce Board conducted a long and deliberative process. The Board established a technical workgroup to develop a proposal for its consideration. The technical workgroup consisted of state and local staff from each of the affected programs, mostly research managers and some local program administrators. The technical workgroup began by identifying the basic goals that the programs were all trying to achieve. After the Workforce Board agreed to the goals, the workgroup proceeded to identify performance measures for each goal. The workgroup identified a list of a couple of dozen measures, a long list, in order to paint a fairly complete picture of performance. Nearly two years after beginning this effort, the Workforce Board adopted the performance accountability system, including the performance measures.

After four years of reporting performance using the measures, the Workforce Board proceeded to set performance targets. This action was spurred by provisions in WIA and Perkins, and by the interest of the Governor in adding incentives and sanctions to the accountability system. The list of measures, however, was too long for target setting, so the technical workgroup went about identifying a small subset of core measures. To assist in this process, the workgroup conducted a

series of focus groups around the state consisting of representatives of a cross section of programs. Local staff from high school career and technical education programs, community and technical colleges, WIA Title I, and other programs sat down together and discussed the best measures to use if the measures were to be applied to each of their programs. Compromises were reached and measures were proposed by each focus group. The state technical workgroup discussed the results of the focus groups and developed its recommendations. The Workforce Board adopted the recommended state “core” indicators and included them in the state’s unified plan.

IPI Performance Measures

Introduction

What are the best performance measures for workforce development if the same measures are applied horizontally and vertically within the system, including programs that are funded mostly by the states and programs that are funded mostly by Congress? This was a central question considered by the six states—Florida, Michigan, Montana, Oregon, Texas, and Washington—which designed this Blueprint. The state teams, with representation from a cross-section of workforce development programs and the assistance of the NGA’s Center for Best Practices and the Ray Marshall Center at the University of Texas, discussed the answer over the course of three national meetings. The states began by considering what policy leaders want to know about performance results. They then suggested the selection criteria for judging measures, and analyzed the advantages and disadvantages of a long list of possible measures. In the end, the states agreed on a relatively short list of measures that best respond to the performance questions commonly posed by policy leaders. This section of the Blueprint summarizes the discussion and recommendations of the six states’ teams. Other states may wish to follow suit, either by following this type of process within their state, or by adopting some or all of the recommended measures.

Before beginning, it is important to note what the performance measures do *not* provide. First, the focus of the Blueprint is on performance information that meets the needs of policy leaders—elected officials, cabinet officers, and program heads—not the needs of day-to-day program managers. The information needs of policy leaders are often different than those of program managers. Program managers and their staff need up-to-date, preferably real-time information that enables them to make service delivery changes and other program decisions and turn things around quickly if they are not going well. They do not need such immediate information to be consistent with the information used by the managers of other programs. Policy leaders, on the other hand, tend to make changes more deliberately. Legislators make statutory changes only once a year, or in many cases, only every other year. The Congress provides oversight on a longer schedule. Education leaders have little need for the latest quarterly report on program placement rates. Policy leaders involved in systemwide policy or planning, however, do need information that is consistent across programs.

There are also methodological reasons for differences in performance information at the policy and management levels. The best sources of data for consistent information are administrative records; not only are these sources inexpensive, they enable one to apply consistent methodology across programs. One can use the same data source and methods, for example, to calculate employment and earnings for community college students and for WIA Title I participants. Other possible data sources, namely case manager notes or surveys, are either unavailable (community colleges do not employ case managers), too expensive (participant surveys for every program of study at every institution would be extraordinarily expensive), or are not consistently reliable and valid (for example, self-reported earnings in response to surveys).

Administrative records, on the other hand, have their own weaknesses. Most notably, UI wage records, the administrative data source for employment and earnings, have a time lag of about six months in data availability. A six month lag is not a problem for those writing a bill for the legislature, it is for those managing WIA Title I contractors. Making the lag worse is the follow-up period required for policy level measures. Policy leaders want to know whether or not a

program is working. One cannot judge whether or not a program is working based on what happens the day after participants leave the door. There must be a follow-up period of sufficient length to give some assurance that observed immediate results will last. Program managers do not have this luxury.

States and local entities have sometimes dealt with this temporal difference by identifying a “dashboard” of “leading indicators” drawn from data available at exit or shortly thereafter. The hope is that the leading indicators will be positively and strongly correlated with longer term measures of lasting results. Florida has instituted such a “dashboard” of “leading indicators” for WIA Title I, referred to as the “Red and Green Report.” The “Red and Green Report” reports outcomes no later than 45 days after the end of each quarter that can be used for case management, performance based contracting, and project management. Florida has tested the connection of the “Red and Green” measures with the longer term WIA core indicators and found that they are positively correlated. The “Red and Green” name refers to the colors in the report that designate workforce development areas that are performing in either the bottom or top quartile on a measure. This graphic display provides a quick, intuitive picture of how areas of the state are performing.

At the other end of the spectrum are measures of the general state of a state’s workforce. These are not performance measures per se. They do not show the results of workforce development programs or even the results of the system as a whole. They are indicators of how well a state’s entire workforce is doing. They count individuals who never participated in a workforce program, as well as those who have. Oregon Shines provides examples of such indicators: the “percentage of Oregon workers employed in a job that pays wages of 150 percent or more of poverty (for a family of 4)” and the “average annual payroll per covered worker.” Such indicators provide policy leaders with a general picture of how well their state is doing.

The performance measures discussed in the remainder of this section fall in between these two levels. (A state may want to implement a tiered system, as in Florida and Oregon with measures at each of the three levels.) The measures discussed here are neither day-to-day management measures nor general indicators of the state’s well being; instead, they are designed to provide policy leaders with information on how well the workforce development system is performing and whether or not the results are getting better over time. Also, most of the measures can be broken down in order to provide information on performance results by program, subpopulation (i.e., gender, race/ethnicity, disability status), and service strategies, and many of the measures are scaleable from the level of individual organizations to the level of the state as a whole.

I. What Do Policy Makers Want to Know About Performance?

As mentioned above, the six states’ teams began their discussion of performance measures by asking themselves, “What do policy makers want to know about performance?” Drawing from their experience with governors, legislators, and boards, the teams identified the following six types of performance information that policy makers want to know and a series of questions that leaders commonly ask:

1. Labor Market Results for Program Participants

Do people get jobs?
What are they paid?
Do the jobs last?

2. Skill Gains

To what extent do education levels increase?
Do participants continue in further education?

3. Participant Satisfaction

To what extent are program participants satisfied?

4. Social Welfare Results for Program Participants

What are the changes in the receipt of social welfare payments?
What are the changes in the receipt of unemployment insurance payments?
What are the changes in poverty rates?
Are we making taxpayers out of tax users?

5. Results for Employers and the Economy

What are the impacts of workforce development programs on the economy?
Are the programs meeting the needs of employers?
Does the supply of newly prepared workers match labor market demand?
Are program participants better prepared for work?
Do people go out-of-state after they are served?

And policy leaders want to know the relevant information by economic sector.

6. Return on Investment

What do programs cost?
What is the return on investment?

In addition to answering these questions for the workforce development system as a whole, policy leaders want the participant information available by subgroup, such as by gender, race/ethnicity, and disability status. They also want information to make comparisons with other states, to judge what service strategies are most effective (for example, job training vs. job search assistance), and to compare programs.

Answering all of these questions is not possible, and in some cases, not even desirable, at least not without a lot of caveats. For example, it is very difficult to determine the impacts of workforce development programs on a state's economy. A state's economy is greatly affected by national economic conditions such as consumer spending, interest rates, international trade, the federal budget, and where the economy is in the business cycle. These conditions, more than the performance of the workforce development system, determine the health of a state's

economy at any given point in time. Over the long haul, the level of education and training of the state's workforce is certainly a major factor affecting a state's economy, but short-term effects are difficult to isolate.

With regard to participant satisfaction, while everyone agrees that knowing the degree of participant satisfaction is a good thing, after much discussion the teams concluded that measuring participant satisfaction is best done at the point of service as a quality improvement tool, rather than as systemwide measure of performance.

Policy makers are often interested in making comparisons, but one must be very careful in comparing the results of different programs. Different programs serve different populations for different purposes. Given these differences, their results should not be same. If one program's results are better than another's, that in itself should not lead to the conclusion that resources should be shifted from one to the other. The program with weaker results may be serving a population with many barriers to employment that nearly everyone would agree a state should serve, and its weaker results may reflect the need for additional resources.

The ability to understand the outcomes of federally funded programs for the nation as a whole as well as the ability to make comparisons among states would be useful to policy leaders. However, to do this would require standardizing both measurement methodology and the techniques used to adjust measures for variations in economic conditions and participant demographics. This standardization does not currently exist. In fact, the current system makes it difficult for Congress to know what it is getting. States vary greatly in exactly how they apply the same measures and small variations in detail can lead to rather large differences in results. While some states have developed regression models to adjust results for economic conditions and demographic characteristics, most have not, and those that do use different techniques. As a consequence, one cannot tell if differences in performance results among states are due to methodological differences, economic or demographic differences, or real differences in how well states actually perform.

Policy makers want to know the "difference" that programs make. To actually measure the "difference", however, requires net impact analyses that compare outcomes for participants receiving services (the treatment group) with outcomes for comparison groups of individuals similar to the participants but who did not receive the services. It is not feasible to conduct net impact analyses on a frequent basis vertically and horizontally across the workforce development system, and many programs do not lend themselves to experimental, random assignment evaluations.

There are steps that states can take to address these issues, but we should recognize at the outset that despite our best efforts we cannot answer all of the questions that policy leaders ask. In some cases, we have to answer the questions we wish they would they ask.

II. What Makes for a Good Performance Measure?

Given the goal of integrated performance information across the workforce development system, the six states' teams developed criteria to judge the quality of possible performance measures.

Other things being equal performance measures are better the extent to which they:

Are outcome measures: Performance measures should be measures of the results for customers, as opposed to process measures or measures of program outputs. As a part of this criterion, performance measures should measure things that programs can substantially affect, not variables determined by external factors unrelated to program performance.

Promote desired results: Because you get what you measure, measures should be carefully designed to promote behavior and results that are consistent with policy goals. As a corollary to this, measures should attempt to minimize unintended consequences.

Are easily explainable to a lay audience: Policy leaders are busy people and have to digest a tremendous amount of varied information. Measures are more useful to policy leaders the extent to which they are understood quickly and easily, the fewer the measures the better.

Create a level playing field among programs and service strategies: Measures should be designed so that they do not create a bias toward one program or strategy or another. For example: a follow-up period linked to the point of registration is biased against lengthier programs and strategies.

Meaningful for each program: The performance measures should provide meaningful information for each program in the system. They should provide programs with information they care about.

Scaleable and Divisible: Measures should be meaningful vertically across the workforce development system—they should be applicable, to the extent possible, to local institutions, workforce development or other regional areas, states, and the nation. Measures should also be divisible—so that results can be reported for subpopulations and service strategies. States should be able to slice and dice the measures to meet a variety of research and analytical needs.

Timely: There should be a minimal gap between the time of service and the time the performance information is reportable.

Methodologically Sound: Measures should satisfy standards for validity and reliability.

Are not easily “gamed”: While there may be no measure that is completely impervious to manipulation, some measures are more susceptible than others. Measures should be constructed to minimize the opportunities to improve results by manipulating the measurement rather than by improving service.

Inexpensive: Performance measures are very important for ensuring taxpayer dollars are wisely used, but states want to minimize the amount of money spent on activities other than direct service to customers, and that includes performance measurement. For the sake of efficiency, measures are better when they rely on data that are already maintained, i.e., administrative records.

Comprehensive and complementary as a whole: The measures should be comprehensive in that, while still a relatively short list, they address all the types of performance information that leaders want. They should be complementary in that if a measure creates a bias in one direction

it should be accompanied by another measure that points in the other direction. For example, if a state uses a measure of earnings that promotes serving people with high pre-program earnings, the state should also use a second earnings measure that promotes serving people with low pre-program earnings.

Finally, it should be understood that there is no perfect set of performance measures. Measures are not capable of perfectly fitting each and every one of these criteria. There are tradeoffs that have to be made. In order to satisfy some criteria, there has to be some give on others.

III. IPI Performance Measures

The six states' teams identified eight performance measures that best meet the above criteria and provide information that policy leaders want to know about the performance of the workforce development system. The following section presents the eight measures and some important "nesting" or sub-measures that states or others could use for informational purposes. The measures are grouped into broad categories that correspond with the identified needs of policy makers. The section also analyzes the advantages and disadvantages of the measures and others that were considered. Appendix __ presents definitions for key terms.

The measures are divided into two broad categories: accountability measures and performance indicators. The accountability measures are appropriate for use in setting targets and applying consequences, while the performance indicators are not recommended for these purposes. The indicators do not sufficiently satisfy the above criteria to be used as accountability measures. For example, the employer market penetration measure is not an outcome measure and can have unintended consequences. The return on investment measures have a long lag between the time of service and the time the results are available, are expensive to measure on a frequent basis, and the methodology is not sufficiently rigorous to accurately measure small changes over time. Despite these and other shortcomings, these measures are still useful as indicators of the general performance of the workforce development system.

Finally, it should be noted that the measures are not drawn from a list of measures from any of the six states. They are the product of wide-ranging discussions and compromises among individuals from a variety of states representing a variety of workforce development programs. The below figure summarizes the measures.

IPI Performance Measures

Category	Measure
Accountability Measures	
Labor Market Results for Program Participants	<p>1. Short-term Employment Rate: The percentage of participants who have exited with employment during the second quarter after exit. (For youth, enrollment in education counts as well as employment.)</p>

Category	Measure
Accountability Measures	
	<p>2. Long-term Employment Rate: The percentage of participants that have exited with employment during the fourth quarter after exit. (For youth, enrollment in education counts as well as employment.)</p> <p>3. Earnings Level: Median earnings during the 2nd quarter after exit among all exiters with any earnings.</p>
Skill Gains	<p>4. Credential completion rate: The percentage of exiters who have completed a certificate, degree, diploma, licensure, or other industry-recognized credential within one year of exit.</p>
Results for Employers and the Economy	<p>5. Repeat Employer Customers: The percentage of employers who are served, directly or indirectly, who return for the same service.</p>
Performance Indicators	
Results for Employers and the Economy	<p>6. Employer Market Penetration: The percentage of all employers who benefit from the services.</p>
Return on Investment	<p>7. Taxpayer Return on Investment: The present value of the net impact on tax revenue and social welfare payments compared to the cost of the services.</p> <p>8. Participant Return on Investment: The present value of the net impact on participant earnings compared to the cost of the services.</p>

Labor Market Results for Program Participants

- 1. Short-term Employment Rate:** The percentage of participants who have exited with employment during the second quarter after exit. Numerator: The number of exiters with any earnings during the second quarter after exit. Denominator: The total number of exiters. (For programs serving youth, the measure should include enrollment in education as well as employment in the numerator, so the measure captures the positive outcome of continued education.)

2. **Long-term Employment Rate:** The percentage of participants that have exited with employment during the fourth quarter after exit. Numerator: The number of exiters with any earnings during the fourth quarter after exit. Denominator: The total number of exiters. (For programs serving youth, the measure should include enrollment in education as well as employment in the numerator.)
3. **Earnings Level:** Median earnings during the second quarter after exit among all exiters with any earnings. (For programs serving youth, individuals enrolled in education during the second quarter after exit should be excluded from the measure, in order to prevent individuals who are employed only part time because they are students—a positive outcome—from depressing the result.)

Nesting Measures: Measures number 1, 2, and 3 can be measured for subpopulations of participants, such as race/ethnicity and public assistance recipients, and for subsets of services. The earnings data can also be disaggregated to show distributional outcomes; for example, the median income among the bottom quintile of exiters, the second quintile, and so forth.)

Discussion: The recommended measures provide simple and straightforward answers to common questions asked by policy leaders. In their simplicity, they avoid many of the disadvantages of more complicated measures sometimes used. Other possible measures include looking at employment only among those unemployed at registration or “pre/post” changes in earnings. Both of those approaches can be misleading. They imply that the services caused the observed change in employment or earnings, as though all individuals not served would have remained unemployed or had stagnant earnings. We know, however, this is not the case. In order to measure the change caused by program services, net impact measures are necessary.

Limiting the employment rate measures to only those unemployed at registration also excludes a substantial share of workforce development participants and makes the measures largely meaningless for important programs. About 70 percent of community and technical college students are employed at registration. To have a key performance measure that excludes the majority of participants in the largest adult training program in the nation does not make sense for integrated performance information.

Pre-post changes in earnings are mostly determined by the pre-program level of earnings. The higher the pre-program level of earnings the lower the gain, other things remaining equal. Unless results are adjusted for the level of pre-program earnings, differences in this measure from year to year are more likely to reflect changes in pre-program earnings than changes in program effectiveness. There are other challenges as well, for example picking the right pre-program time period to use. The first quarter prior to registration participants typically have experienced a decline in earnings. Should the first pre-program quarter be used, perhaps artificially inflating the result, or should an earlier quarter be used, perhaps artificially understating the program’s effectiveness?

The number produced by a measure of pre-post changes in earnings is not something that is intuitively meaningful. While \$30,000 per year is an excellent result for most workforce development services (other than the labor exchange), it is not so obvious whether a pre-

post change of \$10,000 is good or bad. The answer is: it depends on the finishing line. Going from \$0 to \$10,000 is not good since it still does not provide sufficient earnings to live on. Going from \$10,000 to \$20,000 is generally a good outcome, but one cannot tell that by a pre-post measure alone. The intuitive meaning of the number is particularly important if the earnings measure is going to be reported as part of a consumer report system. An individual trying to decide which training program to enroll in is much more likely to want to know how much money students typically make after the program, than to want to know the average pre-post change in earnings.

For the earnings measure, we recommend the median rather than the average (mean) because for many workforce programs only about a third of participants earn at or above the average. The average, therefore, creates a misleading impression about performance and the earnings participants can expect. Although the measure is based upon earnings in one quarter, it might best be presented as an annualized number (quarterly earnings times four) since annual earnings levels are more easily understood. Finally, while a level of earnings measure does not in itself provide a point of comparison to a preferred policy goal, a state (or an individual for that matter) could use whatever is preferred—the poverty threshold, a self-sufficiency standard, a “living wage,” or some concept of a “family wage”

The choice of which follow-up period to use is an important one. While the first quarter after exit provides the most timely information, it is subject to manipulation by timing the date of exit. It is also incomplete. Many participants do not become employed until some time during the first quarter after exit; for these, the first quarter will not capture a full quarter’s worth of earnings. This is most obviously the case for secondary career and technical education. For adults, the longer and more costly the investment in training, the longer participants are willing to wait to find the right position in their field. Their earnings prior to finding that position do not capture the result of the training. There may also be a delay due to the timing of licensure exams. For these reasons, many argue that the third quarter after exit is the first quarter that will fully reflect the earnings outcome. This however, is a long time to wait for information. Consideration must also be given to the criterion of a level playing field. Longitudinal analyses of earnings shows that the longer the follow-up period, the better the relative results for training services compared to quicker job search services. On the other hand, one could make job search look relatively better by using the shortest follow-up period possible, such as the quarter of exit. Given all these considerations, the second quarter is recommended as the best follow-up period. While the timing of the first follow-up period for employment is not as sensitive to these issues as is earnings, we recommend the second quarter for that as well in order to use the same follow-up period for as many measures as possible for the sake of simplicity and ease of explanation.

The preferred measure for a longer term follow-up is the employment rate during the fourth quarter after exit. This is not a “traditional” retention measure. The concept of retention measures fits well with workforce development programs that have a strategy of placing participants and then helping them retain employment. For many programs, including community and technical college job training and secondary career and technical education, this is not part of what they do. Employment retention measures can also be misleading if they measure retention in employment only of those employed during the earlier follow-up period. One could observe an 80 percent employment rate in the first period and an 80 percent retention rate in the second, both of which sound fairly

impressive. But it would be possible that only 64 percent of all exiters could be employed during the second period—a poor result that would not be revealed by a traditional retention measure.

Policy leaders also express interest in measures of earnings change during the post-program period. Put colloquially, many leaders want to know if individuals are stuck in dead-end jobs or if they experience wage progression. Research shows that increases in earnings during the post-program period are strongly correlated with continuous employment. Post-program earnings gains are also correlated with the level of initial earnings—those who land a good job tend to experience the largest subsequent gains. To some extent, therefore, measures of post-program earnings gains are redundant if one already has measures of earnings level and short and long-term employment.

While measures of earnings change and employment retention do not rise to the top of the list of our recommended measures for the reasons cited, they can provide useful information and states may wish to use them for analytical purposes. Possible measures include:

Employment retention: The percentage of exiters who are employed during each of the second, third, fourth, and fifth quarters after exit. Numerator: the number of exiters with any earnings during each of the second, third, fourth, and fifth quarter after exit. Denominator: the number of exiters.

Pre-post Earnings Change: The difference between the level of earnings during the second and third quarters after exit and the level of earnings during the second and third quarters prior to program registration. Numerator: the level of earnings during the second and third quarters after exit minus the level of earnings during the second and third quarters prior to program registration among those with earnings during both of these pre- and post-program periods. Denominator: the number of exiters with earnings during both of these pre- and post-program periods. (The restriction of the measure to those with earnings during both periods is to make it a measure of earnings, as opposed to a measure of the compound effect of earnings and employment rates.)

Post-program earnings change: The difference in the level of earnings during the fifth quarter after exit and the level during the second quarter after exit. Numerator: earnings during the fifth quarter after exit minus earnings during the second quarter after exit among those with earnings during both quarters. Denominator: the number of exiters with earnings during both the second and fifth quarters after exit.

Skill Gains

- 4. Credential completion rate:** The percentage of exiters who have completed a certificate, degree, diploma, licensure, or other industry-recognized credential within one year of exit. Numerator: the number of exiters who have completed a credential within one year of exit. Denominator: the number of exiters.

Nesting Measures: Measure 4 can be measured for subpopulations of participants and for the subset of participants that received training. The measure also incorporates the data to report the number of exiters completing a credential, in addition to reporting the rate.

Discussion: One of the key issues in the design of a skill gains measure is determining what to count as a gain. Experience with WIA Title I shows that unless there is a fairly tight definition, there will not be consistency in how gains are measured across geographic areas and programs. For the sake of consistency and in order to count gains that matter, the recommendation is to count credentials only. This should include standard credentials—such as academic certificates, diplomas and degrees—and also industry-recognized credentials such as licensure and industry-based certifications. Ideally, credentials should be based upon assessments of students against industry standards. This, however, is not yet sufficiently common to make it a requirement in the definition of what constitutes a credential.

One of the technical challenges in measuring credentials is dealing with the large number of individuals going in out of short-term training during the course of their working life. Many community colleges have responded to this type of demand by implementing modular training with multiple exit points. Reaching an exit point signifies the attainment of meaningful skills. It would be useful for the credential measure to positively count this type of completion, particularly when it satisfies industry-based standards.

Another issue is whether to include all exiters in the denominator or only exiters who received training. The recommendation is to use all exiters because it will encourage programs to promote training and skill development for the whole population of participants. If only those exiters who received training were counted in the denominator it would be possible for a state to achieve a credential completion rate of 100 percent while training very few individuals. For those interested in knowing the credential completion rate among only those who receive training, that can be measured as a nesting measure.

Another important point of information is the *number* of individuals who attain new credentials. This is a measure of the extent to which workforce development programs and the system as a whole are increasing the supply of skilled workers. Since this number is the numerator in the credential completion rate measure, it can be readily provided as a nesting measure.

Participant Satisfaction

As mentioned earlier, the state teams came to the conclusion that participant satisfaction is best measured at the point of service as part of quality improvement efforts, and not as a performance measure for the system.

Social Welfare Results for Program Participants

The states' teams decided that social welfare results are best measured through measuring subpopulation results for earnings and through the return on investment measure (number 7, below). The primary policy question regarding social welfare participants is, "How much are they earning after participation?" This can be measured by looking at earnings measure (number 3) for the subpopulation of social-welfare recipients one is interested in.

As part of measure number 7, a state can look at the net impact of workforce development participation on social-welfare payments. This is a better performance measure than the percentage of participants receiving social welfare payments after participation. This latter measure is more related to policy changes in eligibility requirements, than it is to program performance. A net impact analysis enables a state to judge how much difference program services make, since the comparison group will be subject to the same eligibility policies as the participant group.

Results for Employers and the Economy

- 5. Repeat Employer Customers:** The percentage of employers who are served, directly or indirectly, who return for the same service. Numerator: the number of employers who receive the same service again within a year of the initial service quarter. Denominator: the number of employers who receive a service during a quarter.
- 6. Employer Market Penetration:** The percentage of all employers who benefit from the services. Numerator: the number of employers served either directly or indirectly. Denominator: the total number of employers.

Nesting Measures: Measures number 5 and 6 can be measured for particular sectors of the economy or other subsets of employers that are target markets, and for particular services.

Discussion: It is difficult to measure the outcomes of workforce development programs for employers and the economy, especially in cases where employers are not the direct customers but are indirectly served by hiring program participants. Economists can argue that the benefit to employers can be measured in the additional earnings paid to program participants since employers would not pay the additional earnings if former participants did not produce value at least equal to that cost. From a political perspective, however, it is difficult to convince elected officials and business leaders that the result for employers is best measured by increased individual earnings. Another potential measure, productivity gains, is notoriously difficult to measure and is not very scaleable. It is also highly subject to factors external to workforce development programs. Serious consideration was given to a measure of employment retention with the same employer. Retention with the same employer may be a sign of employer satisfaction with the quality of the participants' work. Participants, however, may leave an employer for good reasons (for a better job) as well as for bad reasons (they were fired), and employment retention with the same employer would count both as though they were bad. In fact, participants often do better when they do not retain their employment with the same employer especially if they are beginning their career at the entry level.

The measure of repeat employer customers is recommended as the outcome measure for this category. The measure includes indirect as well as direct service in order to include employers who hire program exiters. An increase in repeat customers over time signifies that workforce development programs are providing services that employers want. The measure can be viewed as a proxy for employer satisfaction. To measure employer satisfaction directly would require surveys that would be expensive to administer vertically across the system.

Employer market penetration is more an output measure than a measure of the outcomes for employers. It is useful information for policy makers to know, although some might question whether higher is always better since it could involve competition with the private sector. The measure refers to employers who “benefit” from services, to reflect that both direct and indirect services should be counted. Including all employers in the denominator will generate a number that is relatively small. States may consider limiting the denominator to employers that hire employees. Job placement, however, is not the only service provided to employers, and limiting the measure to only employers that hire would create a disincentive to market services to all employers. The results of all services to employers are to be counted in the measure, including, but not limited to providing labor market information, training participants who become new hires, and upgrading the skills of incumbent workers.

A technical challenge with both measures is that care must be taken to identify employers. Potential problems include handling employers with multiple locations and employers whose legal identity changes. The measure of repeat employer customers has the potential downside of creating a disincentive to serve small employers who are less likely than large employers to need to repeat the same service. This problem might be mitigated through a regression model or other adjustment methods. Both measures treat small and large employers alike, even though there can be a much larger economic impact from serving a large employer than a small employer. It would be easier to increase ones market share by targeting small employers. This potential unintended consequence is another reason for treating market penetration as a performance indicator rather than accountability measure.

Return on Investment

- 7. Taxpayer Return on Investment:** The present value of the net impact on tax revenue and social welfare payments compared to the cost of the services. Numerator: estimated net impact on tax revenue and social welfare payments to 67 years of age. Denominator: cost of the services.
- 8. Participant Return on Investment:** The present value of the net impact on participant earnings compared to the cost of the services. Numerator: estimated net impact on participant earnings to 67 years of age. Denominator: cost of the services.

Nesting Measures: Measures number 7 and 8 can be measured for subpopulations of participants and for subsets of services.

Discussion: Measure number 7 attempts to show how much taxpayers receive for their investment through increased tax revenue and reduced social welfare payments. Measure number 8 attempts to show whether, and to what degree, the impact on participant earnings is greater than the cost of the services. The two measures might more accurately be termed “cost effectiveness” measures, but they are referred to here as return on the investment measures because that term resonates more with policy leaders. The underlying premise of both measures is that the return to taxpayers and participants is based on the net impact. To count the observed outcomes as if nothing positive would have occurred to individuals if they had not received services is misleading.

The recommended methodology for the return on investment measures is to statistically construct a comparison group. The most practical data source for that purpose, for most programs, is the database of Employment Service registrants. This database contains a rich array of information on each individual—including age, race/ethnicity, gender, disability status, prior education, and location—and can be linked with employment wage files and other data sets to obtain employment and earnings data, and history of unemployment benefits and public assistance receipt. Statistical matching can be used to find the Employment Service registrants that most closely match the workforce development participants of interest. This data source also has the advantage that most of the comparison group share with the participant group the characteristic of being job seekers or incumbent workers who want to improve their situation. While the database of Employment Service registrants is the most useful source for statistically constructing a comparison group, it is not the only source, and for some populations—particularly in-school youth—other sources are needed as well.

Net impacts can be assessed by comparing outcomes for individuals who received workforce development services (other than registration with the Employment Service) to their matched counterparts. Employment Service registrants that ever participated in any other workforce development service are removed from the comparison group pool. This method estimates the impact of workforce development services beyond the impact of simply registering with the Employment Service. Clearly, this method cannot be used to measure the impact of registration in itself.

The impact on tax revenues can be derived by formula from the impact on earnings. State departments of revenue maintain formulas that estimate the amount of tax revenues generated by changes in earnings. The change in tax revenues can, therefore, be estimated without accessing individual tax records and regardless of the types of taxes in a state. Administrative records are available data sources for earnings and social welfare payments. Also, in calculating costs one must be careful to include all costs, which may include federal, state, and local expenditures.

It is also important to include foregone earnings of participants in the calculation, and to discount benefits in order to compare benefits and costs in terms of net present values. In order to measure the full net impact, estimates should be based upon the observed period of time and statistical models of the benefits out to the age of 67 (future eligibility for social security benefits). In order to have sufficient data to estimate benefits to age 67, it is necessary to have several years of observed data. As a result, there is a considerable lag between the time that participants exit service, and the time that results are available. This is one reason why the return on investment measures are not appropriate as performance accountability measures. By the time the results are available, much will have already changed—frequently including the program administrators.

While this methodology does not withstand the scrutiny of those who believe that only an experimental design is sufficient, it does offer a practical and relatively inexpensive way of estimating net impacts. Moreover, it is much more rigorous than the methodology behind most so-called return on investment measures presented to state policy leaders. Washington's Workforce Board has applied this methodology twice to measure the return on investment for workforce development, and has received good reviews from the Office of the Governor and the state Legislature. While the numbers created by the measures

should be considered as only general estimates (another reason they are not suitable as accountability measures) they do give a sense of whether services produce positive net impacts that exceed costs, and the general magnitude of the net impacts.

For programs for which random assignment is appropriate, these state measurements can be complimented by national experimental net impact evaluations. These research-based evaluations satisfy stricter requirements for scientific rigor and provide valuable information regarding the actual impact of programs and strategies, including their effectiveness for various sub-populations. As a package, then, the performance system would include a variety of measurements—output, outcome, and experimental and non-experimental net impact measures—that complement one another.

Data Sources and Supplemental Data

The data sources for these performance measures are administrative records. These records include the following sources for outcome data:

- Unemployment Insurance wage records
- Department of Defense personnel records
- Federal Office of Personnel Management records
- United States Postal Service personnel records
- College and school student records
- The National Student Loan Clearinghouse

The major gaps in these administrative records are out-of-state employment and data on self-employment. WRIS provides out-of-state employment information for services authorized by WIA, however, the scope of WRIS must be expanded to other programs in order to fill the first gap for out-of-state employment wage records (see p. ____). DOL's FEDES project is developing a national system for accessing Defense and federal personnel records. States may now individually reach agreement with one another and the federal agencies for access to employment records. It is not practical, however, for 50 states to be doing this individually. These national systems need to be accessible for the full range of workforce development programs.

For self-employment there are three major options, each with serious flaws. The first option is to use state tax records. These records, however, often do not have quarterly data, and for those states without an income tax, they do not include earnings. The second option is, in cases of self-employment, to allow supplemental data, such as case manager follow-ups with documentation. This option, however, is expensive and will be difficult or impossible to apply evenly across programs, areas, and states. The third option is to not count self-employment. This option, however, creates a disincentive to offering entrepreneurial training and other services related to self-employment. There is no perfect solution to this dilemma, and self-employment may be increasing as firms contract out more work.

The six states' teams recommend that state integrated performance information be based upon administrative records, but that individual programs, particularly at the local level should include supplemental data for their own use as meets their needs. This dual approach supports consistency but recognizes that more complete information may be important for management purposes. The supplemental data should be documented and verifiable, and subject to periodic audit.

A final important point on supplemental data: whatever data sources are used to measure performance should be the ones used to develop performance targets. It hurts credibility to compare results against targets based on different data. For example, one would not set targets based upon data from administrative records and then allow programs to include supplemental data that will artificially boost results, perhaps above targets.

IV. Performance Targets and Consequences

Having designed performance measures and measured the results, the next step is to compare the results to something. As noted earlier, states should not use the performance measures to simplistically compare programs with each other. Different programs serve different populations for different purposes and their results should not be expected to be the same. Over time as there is longitudinal information, an obvious comparison to make is with a program's own past performance in order to judge whether performance is getting better. To the extent that there are consistent performance measures among states, another likely comparison is with other states, perhaps benchmarking against states with best results. The third obvious comparison to make is with performance targets.

Historical data is needed in order to set performance targets. A state must examine how it has performed in the past in order to know how far it should go in the future. Once there is sufficient baseline data, a state can set performance targets, most likely in incremental steps above past performance. Targets for incremental improvement are not, however, always logical. If a state is already performing at or near 100 percent, or among the best in the nation on a measure, then it makes sense to focus improvement on other measures. One way of doing this, once there is comparable measures among states, is to set targets for incremental improvement except for measures where a state is already in the top quartile among the states.

To be clear, states should not set targets at the same level of all programs. Targets for each program should be set at a level appropriate for that program. The earnings target for a program that serves youth, for example, should not be the same as the earnings target for a program that serves dislocated workers.

What should be the consequences for whether or not targets are met? The original IPI state teams were of mixed views on this issue. They were, however, in agreement that whatever types of consequences a state chooses, they should be phased-in over time from the least significant consequences to the most significant in order to allow for a period of adjustment. This is important for stakeholder buy-in and to work out kinks in the system, including the accuracy of performance reporting. (As the consequences become more significant, staff are more likely to be responsive to correcting errors in performance measurement.) It is also important to use a collaborative process among all the affected parties to develop the system of consequences.

Simply reporting results and targets in the same performance report has consequences. It is likely, for example, to influence policy leaders who see the information when they make statutory or budgetary decisions. Perhaps the next level in a continuum of consequences is non-monetary recognition. Done correctly, this can be a significant motivating factor. A third level of consequences for positive performance is financial. With regard to sanctions, a common first step when performance fails to meet targets is requiring improvement plans. The most severe steps are financial sanctions and a reorganization or replacement of the responsible authority.

A dilemma, however, is whether funding should be reduced for poor performance or if poor performance should be taken as a sign that more funding is needed. Also, if funding is reduced for poor performance, this may make it harder to improve performance since that often takes additional resources to accomplish. Further, there is concern that reduced funding harms the community—or places it at an additional disadvantage. There is also a concern that the more performance is linked to financial consequences, the more there is an incentive to “game” the numbers. While the IPI measures are designed to minimize the opportunities for “gaming,” the possibility cannot be eliminated. One way of minimizing these issues is for a state to place restrictions on funding rather than to reduce funding for poor performance. The most common restriction is to require that funds be used to finance an improvement plan approved by the state. This approach lowers the stakes that can lead to “gaming” and provides resources that are directed to improving performance.

Not all performance measures are appropriate for financial incentives and sanctions. *Among the IPI measures, we recommend that financial consequences not be linked to the measure of employer market penetration and the two measures of return on investment.* These three measures are better used for informational purposes, rather than as accountability measures linked to specific financial consequences. Policy leaders will still see the information and, as noted earlier, the information will likely influence budgetary decisions.

Florida has perhaps the most elaborate state system of performance targets and consequences linked to performance. For example, the state sets performance targets for each local recipient of Perkins (federal career and technical education) funds. The targets are based upon incremental improvement from past performance. For each measure, the performance of local recipients is divided into quartiles. The target for those performing in the lowest quartile on a measure is set two percent above past performance, for those in the second quartile the target is 1.5 percent above past performance, for those in the third quartile the target is 1 percent above past performance, and for those in the top quartile the target is to maintain past performance. Variations from this formula can be negotiated for such events as major economic dislocations. For any measure where improvement is needed, local recipients must submit a brief plan for enhancing performance.

Florida has also established an incentive pool for regional workforce boards using federal funding from WIA section 503 awards and Temporary Assistance to Needy Families (TANF) High-Performance Bonuses. All regional workforce boards that achieve their individually negotiated performance targets for WIA Title I, Wagner-Peyser, and TANF are awarded minimal performance incentives. The regional workforce boards that perform in the top quartile compared to their peers get a larger performance award. The state also awards high performing boards with public recognition and plaques at an annual Workforce Summit. The state requires low performing regional boards to provide a corrective action plan. If low performance persists, they must have a state-approved program improvement plan, often supported by the state and

peer technical assistance and sometimes funding for specific interventions. Continued poor performance results in progressive levels of on-site oversight by state staff or state-designated peer supervision, “receivership,” or replacement of local executive staff, all of which have happened. So far no region has been redesignated or consolidated due to performance problems, but that option remains as an ultimate sanction.

Adjustments

Performance measures and targets have the potential downsides of creating disincentives to serve “hard to serve” customers, and rewarding or sanctioning programs for results caused by external factors. Advocates for hard to serve individuals are rightfully concerned that programs might, as a result of performance measures, avoid serving the people they represent. Program staff are justifiably concerned they might be punished if an economic downturn hampers their results. Regression adjustments are one tool that states can apply to minimize these potential pitfalls, improving the validity of the measures and stakeholder acceptance. Regression adjustments “untangle” program performance from the effects of labor markets and personal characteristics. States can develop mathematical regression models to statistically control for changes in economic conditions and the demographics of program participants. States can apply these models to adjust either targets or results.

Washington’s Workforce Board has developed regression models for the federal core indicators for WIA, Title I, Perkins, and Adult Education and Family Literacy, and for its state additional measures. (See Appendix ___ for more information.) The Board has successfully used the models to negotiate revisions in targets with DOL and DOE. The Workforce Board has also applied the models to adjust local targets prior to allocating incentive dollars to local areas. In order to be logically consistent, the Workforce Board adjusts targets both upward and downward depending on which direction the models indicate. So for example, if an area’s participants have included fewer people with disabilities, fewer homeless people, and fewer people who are high school dropouts than the baseline period used to establish the targets, the targets for that area are raised. In June, 2004 the Workforce Board once again requested that DOL agree to revised state targets for WIA Title I, including for the first time upward revisions on some state targets. DOL agreed to the targets proposed by the state.

At the present time, Michigan and its contractor, the W. E. Upjohn Institute for Employment Research, are developing regression models under the sponsorship of DOL. Upjohn is designing the models for the common performance measures developed by the OMB and adopted by DOL. Innovations of the Upjohn approach include adding real-time leading indicators (such as employment at exit), economic forecasts, and participant demographics at registration, so that the models estimate the probable impacts of changes in economic conditions and participant demographics before the results are available, and even before the actual economic conditions and demographics of exiters are known. Upjohn is designing the models with local managers in mind. The goal is to provide a package whereby a manager could check how his or her site is currently performing against targets given participant characteristics and the economic forecast. In contrast, the Washington models are based upon observed data only. Revisions in targets are made only after data demonstrates that actual economic conditions and exiter demographics have changed.

The Upjohn models are not yet final, but so far Upjohn is discovering that prior earnings and employment of program participants are very important variables in determining results, particularly for the OMB measures of entered employment and pre-post earnings gains. This finding goes back to an earlier discussion in this Blueprint, that measures of pre-post changes are highly subject to the pre-program characteristics of participants, and heightens the importance of using adjustment methods if such measures are to be useful.

A major issue is whether there should be national, state, or perhaps regional models. Upjohn argues that state models have the advantage of access to individual unit data that are not available at the national level, and are not affected by unmeasured differences among states. Unmeasured differences among regions of the nation lead some to advocate for regional models. National models, however, would enforce methodological consistency and would therefore create a more level playing field among the states. This would be useful given the current wide variation in targets between federal regions.

A final issue is whether targets should be adjusted solely based upon mathematical models, or if other factors should be allowed into negotiations. The Workforce Board follows the practice of basing adjustments on the models, but negotiates further adjustments when hard data is presented that the models missed. For example, a workforce development area successfully argued that the WIA administrative data on incidence of youth incarceration were incorrect by presenting the results of a data match between the WIA records and county court records. Apparently youth participants were not fully forthcoming in self-reporting their history of incarceration to their case managers, who in turn entered the incorrect data into the administrative records. Based upon that information, the state lowered the area's targets for youth.

Shared Information Systems¹

Introduction

A major barrier to integrated performance information is the absence of a shared or integrated information system for workforce development in most states. Participant information remains scattered about in various program level MISs, and there is no system for integrating the information from the multiple MISs to support performance measurement and reporting.

Efforts to create an integrated workforce development database can follow one of two major paths. The first, and more ambitious approach, is to replace existing MISs with a new system designed to meet the operational and performance analysis and reporting needs of the workforce development programs within a state. Developing such a system would be a massive and expensive undertaking and would require workforce agencies to abandon their legacy data systems and embrace the new, shared system. None of the six states has followed this path. The second approach is to develop a “data warehouse” that is built on top of existing MISs. This approach does not replace existing MISs that are still necessary for program management. The warehouse is a longer term repository where data are cleaned and matched in order to analyze and report performance outcomes.

A data warehouse has four features that distinguish it from an operational MIS. First, a data warehouse is oriented around subjects. While an operational system is designed around functional activities (inputting applications, training activities, fiscal information, etc.), a data warehouse is organized around major subjects (customers, Workforce Investment Boards, providers, etc.). This subject orientation presents the data in a much easier-to-understand format and facilitates analysis of data for strategic planning.

The second feature of a data warehouse is that it is *integrated*. Integration of data within a warehouse is accomplished by making the data collected from various sources consistent in format, naming, and other aspects. For example, a program MIS may represent “male” and “female” differently by using coding schemes such as “m” and “f,” by “1” and “2,” or by “b” and “g.” Often, the inconsistencies are more complex and subtle. In a data warehouse, such differences are resolved so that the information is always maintained in a consistent fashion.

A data warehouse maintains both historical and more recent data. An operational database, on the other hand, maintains the most current, up-to-date data. A data warehouse contains data that is generally loaded from the operational databases on a periodic basis, and maintained for a period of three to ten years. This is a major difference between the two types of information systems and of extreme importance to decision makers, who want to understand trends and relationships between data.

A final characteristic of a data warehouse is that the information is accurate in reference to a given point in time. A warehouse continually absorbs new data and integrates it with the previous data. Once the data is loaded into the data warehouse, it remains unchanged as much as possible, although changes can be made, for example, to accommodate new definitions or business rules.

¹ This section of the Blueprint borrows heavily from the paper, “Developing a Data Warehouse for Workforce Programs,” by John Baj, The Center for Governmental Studies, Northern Illinois University, March 2004.

It should be clear that a data warehouse and an operational system are orientated toward different purposes. If the desire is to develop a resource so that caseworkers and frontline staff can examine the current status of customers across various programs, then a multiple program MIS should be developed. If the desire, however, is to create an executive level performance information system designed to examine historical trends and support strategic planning, a data warehouse is the more appropriate choice. A state can also connect a data warehouse to more current outside data, through Operational Data Stores, allowing analysts, for example, to compare recent performance with historical data. While a data warehouse can be designed primarily to provide policy level information, it can also be a rich source of data for program managers.

This section of the Blueprint discusses the key issues and decisions faced in the establishment of a data warehouse for workforce development programs. Many of the issues are interrelated, and some decisions restrict the range of options that are available for subsequent decisions. For example, decisions about the assignment of administrative responsibilities can impact funding options. It is important to think about such interrelationships when considering various options.

Most of the space in this section is devoted to non-technical issues. These are generally the most difficult issues, since the technology has improved to the point that states generally have staff or contractors available who can readily handle the technical issues.

I. Authorization and Governance

Perhaps the most important decisions concern authorization and governance. If system level data collection and reporting is to become an ongoing activity in a state, the data warehouse should be formally authorized to undertake these activities. Authorization invests the data warehouse with the legal authority to conduct its operations.

The authorizing, or charter document may take one of three basic forms: legislation, executive order, or a binding agreement among participating agencies. A state may also use a combination of these approaches.

In Texas, legislation mandated the creation of the Automated Student and Adult Learner Follow-Up System. The legislation states, “The follow-up and evaluation system shall be used to assist the council, local workforce development boards, institution boards, the Texas Higher Education Coordinating Board, the Central Education Agency, and other agencies in evaluating the labor market success and effectiveness of workforce development in this state.” The legislation goes on to say, “Each state agency represented on the council shall provide information to support the Council’s follow-up and evaluation system as requested.” (Senate Bill 642, 1993)

In Washington, legislation directed the Workforce Board to set standards for data collection and maintenance and to develop a database with records on enrollments, program activities, and job placements. The Workforce Board subsequently joined with the Employment Security Department, the State Board for Community and Technical Colleges (SBCTC), and the Office of Superintendent of Public Instruction to adopt an interagency agreement establishing an interagency consortium to authorize data linking for the purpose of measuring participant outcomes. This dual approach of general legislative authority and interagency agreement offers legitimacy and agency buy-in, plus the flexibility to make changes over time without the need to return to the legislative process.

II. Purpose of the Data Warehouse

It is very important for the purpose of the data warehouse to be clearly stated in the authorizing document. Typically, the statement of purpose is a broad declaration that is further refined by the governing authority. It should convey the primary mission of the warehouse and help the governing authority define priorities. The statement can be crafted to define the operational boundaries of the data warehouse. For example, the authorization can address the question of whether the data warehouse will be confined to systemwide reporting or whether it will also be responsible for sub-state or program level reporting. In Washington, the purpose statement includes, "Manage a cost-effective process of compiling program outcomes data for use in program improvement activities at the local and state level, and for use in evaluations thereof, as needed for reporting to the legislature, agency management, and to consumers."

III. Leadership and Oversight

The charter document for the data warehouse should identify the entity that will be responsible for leadership and oversight. These responsibilities include defining and establishing the policies regulating the use, access, and reporting of information. In addition, the governing authority will be responsible for making sure the data warehouse is fulfilling its mission in a timely and cost effective manner.

Governance for the data warehouse can be vested in a number of alternative entities. The four major options are the state's Workforce Investment Board, a single state agency, an interagency committee or council, and the Governor's office.

A state Workforce Investment Board has coordinating and evaluation responsibilities and representation across programs, these features make it a reasonable entity to serve as the governing body for the data warehouse. On the other hand, the governing authority for the data warehouse will be responsible for regulating uses of and access to the database and these may be more operational roles than some state Workforce Investment Boards are used to performing.

A second option is a single state agency. While this option may be attractive due to its simplicity, it raises a host of questions. It places one agency in an advantaged position in determining the data collection and reporting procedures affecting other agencies, and access to other agencies' data. Investing leadership and oversight responsibilities of the data warehouse in a single agency may undermine the development of a "systems" orientation since the remaining agencies would have a secondary role in decision-making.

The third major option is to have the data warehouse governed by an interagency entity, whether a council, consortium, or committee. The interagency entity could include representation from each of the agencies contributing participant records to the warehouse, and also agencies contributing outcome data, such as employment wage records and public assistance data. The major advantage of this approach is that each agency maintains a greater degree of control over the use of their own data than with the other options. A major disadvantage is that an interagency entity will have the normal challenges of a participatory process that tend to slow down decision-making.

A fourth option is to place the data warehouse under the direct authority of the Office of the governor. This alternative will give the office of the governor greater control and access to information on the state's workforce development efforts. It may, however, result in instability in the system as administrations and staff change over time. Furthermore, the agencies participating in the data warehouse are likely to include agencies that are not under the governor's direct control. Agencies that are under the control of an independent board or a separately elected state official may object to this organizational structure.

What entity will be assigned the leadership and oversight responsibilities for the data warehouse?
State Workforce Investment Board
Single State Agency
Interagency entity
Governor's Office
Combination of above

IV. Funding

While the charter document does not necessarily have to specify how the data warehouse will be funded, it is useful to clarify this from the start. This brings up the question, "How much will it cost?" The answer depends on the functions assigned to the warehouse. Assuming the most limited role of providing annual or biennial systemwide reports on performance, a best guess estimate is sufficient funds to support three positions plus computer time and equipment, and normal indirect expenses. The four basic options for funding are: federal funds, state general funds, agency contributions, and user fees. Again, states can combine these options.

Developing and maintaining a data warehouse for the purpose of strategic planning, program improvement, and performance reporting is a permissible use of several pots of federal funds. For example, WIA Title-I state activities (10 percent) funds may be used. Earmarked state general funds have the advantages and disadvantages of stimulating greater legislative interest. Legislators can become champions of the warehouse, or attempt to micromanage it in unproductive ways. Participating agencies can contribute to a data warehouse operating fund. Funding from multiple agencies may, depending on the state, produce a more stable funding base than a line item in the state budget. If participating agencies fund the data warehouse, there must be some criteria for determining the size of each agency's contribution. A possible criterion is the number of participant records included in the system. A potential drawback of this approach is that the agencies contributing more may feel they should have more influence over the data warehouse. A final source is user fees for the production of ad hoc reports.

Funding options
Federal Funds
State General Funds
Multiple Agency Contributions
User Fees
Combination of above

V. Management and Operation

The entity managing the data warehouse may be the same as the one assigned the leadership and oversight responsibility, but the two may also be different. The major options here include: the state Workforce Investment Board, another state agency, a quasi-governmental agency, or a third party contractor.

One advantage of having the state Workforce Investment Board manage the warehouse is that it would give the state board direct access to the information. Also, by managing the warehouse the state board would gain a greater appreciation of the limits of the data and of the data differences among the programs. A major drawback is that few state boards are currently staffed to do this work. In addition, the activity may distract the state board's attention away from broad policy issues and toward the more technical concerns of data collection and maintenance.

There are likely to be several agencies in a state that have the capacity and experience to manage a data warehouse. Assigning warehouse management to one of these agencies would likely reduce the time and costs of implementing the system. There are, however potential drawbacks. It places data from one agency under the physical control of another, which may be unsettling and in some cases, may be barred by law. There is also the potential problem of setting work priorities if staff are asked to do assignments for the host agency and other agencies at the same time.

Data warehouse management and operation can also be assigned to a quasi-governmental agency or to a third party contractor, such as a university or private business. Both of these options remove the management responsibilities from the contributing agencies and place them in an independent entity. Under these options, data security and access would have to be rigidly controlled through a contractual agreement. These issues must be resolved prior to the transmittal of data, which can delay the initial startup.

In Washington, the consortium of agencies overseeing data matching has contracted at different times with a private business and a state agency (SBCTC). In Florida, a separate unit of DOE manages the warehouse. In Oregon, the matching is conducted by the employment security agency. In Texas, the process is managed by the Career Development Resources Division of the Texas Workforce Commission, formerly the State Occupational Information Coordinating Committee.

What entity will be manage and operate the data warehouse?
State Workforce Investment Board
Single State Agency
Quasi-governmental Agency
Third Party Contractor

VI. Participating Agencies/Programs

The charter document for the data warehouse should specify the initial participating programs and the process for adding or subtracting programs. States may consider whether programs other than workforce development programs should be included. Examples include four-year colleges and universities, workers' compensation, and public assistance programs. There should also be some means by which programs can enter or leave the system as meets their needs. This flexibility is obviously more complicated if the participating agencies are specified by statute only. In Washington, the initial participating agencies were signified by the agencies signing the agreement. The agreement stipulates that additional state agencies may join, but only if agreed to unanimously by the current members. Any agency may withdrawal from the agreement at any time.

VII. Data Sharing Agreements

Another element that could be addressed in the authorizing document is data sharing agreements. Either the governing or administrative entity could maintain the agreements. There could be separate agreements with each of the agencies or there could be an "umbrella" agreement covering data warehouse operations. The former approach might be time consuming and result in a set of agreements with somewhat different use and access restrictions. An "umbrella" agreement signed by all agencies could be more efficient to operate under, but might be more difficult to establish if the agencies have widely different confidentiality concerns.

Data sharing agreements indicate when an agency's data may be shared with another agency. The agreement could, for example, specify that such matching may occur for aggregate statistical analysis but not in cases where an individual's identify could be revealed. Agencies may want the agreement to establish a process for informing each agency when its data is used and by whom.

VIII. "Ownership" of the Merged Database

One of the crucial questions is who "owns" the merged database in the data warehouse. This issue underlies many of the decisions regarding outside access and use of the information. The merged database could be considered a consolidation of the agency databases with ownership of the component elements remaining with the contributing agency. In Florida, data ownership remains with the agency collecting the data. The data is "loaned" to FETPIP in order to create follow-up data from linking records.

The alternative approach is to view the merged database as something more than the sum of its component parts. Under this perspective, ownership would reside with the governing entity and that entity could make independent decisions regarding the access and use of the information.

IX. Data Access

Who will have direct access to the warehouse database and what types of access will be allowed? Some participating agencies are likely to want direct access in order to perform their own research and analyses. An advantage of allowing such access is that it would enhance the value of the database to the individual agencies. Before this access can be given, however, appropriate data sharing agreements must be forged and procedures must be developed to protect the confidentiality of the information.

Assuming that the data sharing and confidentiality issues can be resolved, the next question is what type of access will be given. One option is to allow full access to the entire database. This would enable an agency, if it wanted to, to compare the performance of its programs with those of other agencies. A second option is to limit agency access to information on its participants supplemented with outcome data from other agencies.

There is also the question of whether the data should be made available to external entities. The answer will depend in part on a state's public disclosure statutes. The information stored in the merged database may be considered disclosable public information. Given this, there should be some consideration of developing a public use database. In a public use database, information that can be used to identify specific individuals is suppressed or purged from the database. Obviously, this implies that participant social security numbers are either excluded from the database or encrypted. It might also be necessary to suppress or group geographic/program identifiers that have small numbers to further safeguard the identify of program participants.

Data Access
Will agencies have access to the warehouse database?
What type of access should be allowed?
Will external entities have access to the database?

X. Confidentiality

Confidentiality is an issue that can sink efforts to create or maintain a data warehouse. Recognizing this, Washington State placed first in the purpose statements of the charter document, to "Provide leadership in protecting the privacy and rights of individuals and entities (including firms) whose data records are linked for the purposes of program improvement and evaluation, and to insure the confidential treatment of that information." This language has been in place since 1998, and there have no incidences of inappropriate release of the information.

There should be specific policies and procedures to protect the confidentiality of individual information; for example, the encryption of social security numbers and the storage of data in a secured location. If the data warehouse prepares extracts for use by external parties such as academic researchers, they should be required to take steps to protect the confidentiality of the information.

In the Florida FETPIP system, the data file used for matching contains only social security numbers. Files that contain demographic or program data are retained in a secure environment so that those who obtain the social security numbers cannot connect the numbers with individuals. The files with individually identifiable data are kept under lock and key when they are not in use. Files stored in mainframe computers or personal computer hard drives are subject to stringent security and are accessible only to authorized personnel using secure passwords. The FETPIP system avoids reporting data in individually identifiable form or in groupings where individual identifies could be derived. The data is for aggregate reporting and statistical analysis. It is not used to make decisions about individuals, such as benefit eligibility. Were it do so, a whole another set of legal restrictions would kick in restricting the availability of the source data.

The informed consent of participants may be required for their data to be included in the merged database. The Oregon Attorney General has concluded that the Shared Information System requires the informed consent of program participants. As a result, Oregon asks all workforce development participants to sign an informed consent agreement giving the State permission to use the data for research and evaluation purposes. Other states, such as Texas, feel that this procedure is unnecessary as long as the data is released only in aggregate form. They maintain that this aggregation process is sufficient to ensure the confidentiality of the information.

XI. Information Reporting

What types of reports will be produced using the data warehouse? An expected use is reports on the system-wide performance of workforce development programs. These system reports could simply describe the performance of the system or be analytical reports that relate performance to policy, service strategy, or demographic factors. The data warehouse could also produce reports that show the performance of individual programs. If the reports enable readers to make comparisons among programs, comparisons will need to be conducted in a manner that is agreeable to the participating programs. This is a very sensitive issue that, if handled poorly, could lead programs to withdraw from the data warehouse.

The data warehouse may also be charged with producing ad hoc reports on specific research topics. These reports could be of a technical nature, such as a report on the incidence of missing or invalid data elements, or they could focus on substantive issues, such as a report on participation and outcomes for a population subgroup.

If ad hoc reporting is a role for the data warehouse, a number of questions must be addressed. First and foremost, what guidelines will there be for assessing information requests? What discretion will the administrative entity have to itself identify and explore issues with the database? Will such discretion be limited to technical issues or will it be permitted to explore policy issues? Will the data warehouse respond to information requests from parties other than the participating agencies, the governor's office and the legislature? Will university-based researchers be permitted access to the database? Will staff generate reports requested by such researchers?

If ad hoc requests are going to be accepted, there needs to be a process to evaluate and approve the requests. The following questions will need to be considered: (1) What will be the process for submitting requests? (2) What is the obligation to respond to such requests? (3) What will be the composition of the review committee, if any? (4) How often will the committee meet to review requests? (5) What criteria will be used to evaluate requests? (6) How will priorities be set in cases where multiple requests pass the approval process? (7) How will ad hoc reports be funded? (8) Should such reports be considered a service offered by the data warehouse and accounted for in its annual budget or should they be paid for by the entity requesting the report?

Reporting Issues
What type of System Level Reporting (descriptive versus analytical)?
Will the data warehouse produce individual program reports?
Will the data warehouse produce ad-hoc reports?
What entities can request ad-hoc reports and what will be the approval process?
What will be the guidelines for assessing and prioritizing requests?
How will these reports be funded?

XII. Editorial/Review Policies

There should be some sort of interagency process for reviewing reports that cover more than one program or agency before they are released. Standing policies could be developed that outline permissible content and format as well as technical features such as the method of documentation, and protections for individual confidentiality. Interagency reviews should include both technical and policy issues.

Editorial/Review Policies
What will be the editorial policy for reports?
How will the content of the report(s) be determined?
What review procedures will be used?
What reporting restrictions will be imposed (small N, geography, unit of analysis)?

XIII. System Modifications

As the system matures and workforce development programs change over time, there will be a need to update and modify the data warehouse. System enhancements can fall into one of two categories: the modification and improvement of current activities, and expanding the capacity of the system. It is important for the system to be sufficiently flexible to accommodate change.

System Modification
What will be the procedures for updating the data warehouse?
How will system enhancements be initiated?
What procedures will there be for modifying measures and operations?

XIV. Managing Information Flows

Effectively managing information flows is the key to producing timely reports. One of the first questions is how often the data should be transmitted to the warehouse. This, in turn, depends on how often reports will be required of the system and how often “new” data will be available. For some programs this will be quarterly, for others, this will be annually. Florida’s FETPIP system produces comprehensive data matching on an annual basis, but conducts a subset of matches quarterly.

Data warehouse are dependent on agencies submitting data in a timely manner. Given the many internal demands that are placed on agency MIS units, it may be tempting for them to postpone producing the data warehouse extract. To combat this tendency, it may be necessary to implement a system of incentives or sanctions that can be used to promote the timely submittal of the data files.

There is also the question of how to share the data processing burden. By adopting a pure “data warehouse” approach, agencies would submit data to the warehouse in the form that it is stored on their MIS systems. It would be the data warehouse’s role to resolve the differences in coding and formats to produce the merged database. Advocates of this approach argue that it minimizes the burden on data processing staff and provides greater flexibility in the construction of measures. This is the approach used by Florida’s FETPIP system—while there are several common data elements among the participating organizations, it is important to note that FETPIP imposes no standard definitions or submission formats.

The alternative approach is to require the agencies to take on more of the data processing burden. The entity administering the data warehouse would define a common coding scheme and request the data in a predefined format. The agencies would be responsible for submitting data that corresponds to these parameters. Proponents of this approach argue that agencies have a greater familiarity with their data and are, therefore, better equipped to perform the necessary manipulations.

There must be procedures for agencies to inform the data warehouse of relevant changes in their MIS. Each of the participating agencies will periodically change its MIS format and coding. The data warehouse's computer programs must be prepared to handle these changes. If the agencies do not communicate the changes in advance, the warehouse will not be able to process their data and reporting will be delayed.

There are several alternative approaches for establishing this line of communication. For example, it could be left to the agencies to notify the warehouse of MIS modification on an ad hoc basis. The problem with this approach is that the agencies may not remember to notify the data warehouse due to their intense involvement with their own activities. A second option is for the data warehouse to survey the participating agencies a few months prior to the deadline for the submission of data files. If an agency indicates there have been some changes to their system, the data warehouse could mount a more detailed inquiry to determine if and how those changes will affect warehouse operations. Finally, agencies could include warehouse staff on agency MIS redesign teams, and data warehouse staff could learn about MIS changes as part of being team members.

Managing Information Flows
How often should data be transmitted to the system?
What mechanisms will there be to insure that data is submitted on time?
How should the processing burden be distributed between the data warehouse and the participating agencies?
How will changes in agency MISs be communicated to the data warehouse?

XV. Data Documentation

The information contained in the merged database will originate from various sources. While it is clear that the data warehouse should document the data elements that reside in its database, the level of documentation is less clear. For example, is it sufficient to list the source for each data element and document the manipulations the warehouse performed, or is it necessary to document the procedures used by each agency to extract data from its databases?

Arguments can be made for both approaches. It may be useful to maintain the documentation on data extraction with both the data warehouse and the source agency in order to prevent disruptions caused by a change in staff and a loss of documentation at one place or the other. On the other hand, agencies may be uncomfortable with providing such documentation for fear it may jeopardize the security of their internal systems.

XVI. Quality Assurance

The value of the data warehouse is dependent on the accuracy of the data. It is important that the administrative entity implement quality assurance practices to continually monitor, evaluate, and improve the quality of the database.

To what degree should the data warehouse evaluate the quality of the data supplied by the agencies? Clearly the data warehouse should perform some edit checks, such as procedures to guarantee valid birth dates. For example, the first step in Florida's FETPIP system of processing data is to screen social security numbers to assure that each one meets assignment parameters used by the Social Security Administration and that each number has in fact been assigned.

It is optional whether the warehouse should be responsible for further assessments of data quality. Should the data warehouse be responsible for the physical "cleaning" of the data or will that be the responsibility of the agencies? For example, suppose a problem is found with the data. One alternative would be to have the error noted and corrected in only the warehouse database. Another option would be for the source agency to correct the problem and resubmit a second data file. While the latter approach may be viewed as the preferred course of action, it may delay the process.

A third issue is whether there should be a formal process through which questions regarding data quality can be raised by the data warehouse and addressed by the agencies. A formal approach could include the development of a "problem tracking sheet" used to report the problem and, eventually, document the solution. Although this approach may delay the resolution of data quality issues, it would more fully document the development of the database.

Finally, there is the issue of developing data quality standards. For example, some agencies' databases contain data collected through surveys. Should the surveys have to meet minimal standards for survey methodology in order for the data to be included in the merged database?

Quality Assurance
Should the data warehouse evaluate the quality of agency data or will that be an agency responsibility?
What efforts should be made by the data warehouse to "clean" the data or is that an agency responsibility?
What edit checks will be instituted?
How will the agencies respond to data quality questions raised by the data warehouse?
What degree of influence will the data warehouse have over the data collection activities of member agencies that impact data quality?

XVII. Data Transmittal, Storage and Archiving

The data warehouse may permit agencies to submit data using any format or media they wish in order to minimize the burden on data processing staff. While most of the data is likely to arrive as raw ASCII files, some agencies may choose to submit their files as spreadsheet, ORACLE, or other databases. In addition, a wide range of transfer media could be used by the agencies including tape cartridges, floppy disks and through the Internet. If this “no holds barred” approach is adopted, the data warehouse will need to secure the necessary hardware and software to accommodate these different formats and transfer media. In Florida, the FETPIP system receives data through about every imaginable media.

Data warehouse operations and documentation efforts would be simplified if the agencies were required to submit data in the same type of structure and format. For example, the agencies could be required to submit their data in a fixed field ASCII file that can be read by virtually all software packages. This would eliminate the need to purchase the translation software and to document the data translation process. On the other hand, requiring agencies to produce the data into a specific format could substantially increase their data processing costs and increase the burden on data processing staff.

There must also be a decision on how to structure the merged database. The three major options are a flat, hierarchical, or relational database. Each of these alternatives has its advantages and disadvantages. For example, flat files, which are the simplest type of database, are the easiest to use with standard statistical software packages such as SAS. This makes ad-hoc reporting fairly easy for even the novice user. However, flat files waste storage and are difficult to update when there are additions or changes to data elements. Relational databases, which are the most sophisticated type of database, are very efficient in terms of storage and can easily accommodate changes to the data structure. Ad hoc reporting, however, requires a higher level of user knowledge and expertise.

Data archiving procedures will also have to be established for the merged database. Each year, the data warehouse will add an additional year of participant and outcome information to its holdings. While the information from previous years may be used on occasion to conduct trend analyses, it is unlikely that this information will be accessed often enough to justify online storage and maintenance costs. As a result, some archiving of the information will probably make sense.

Archiving raises a number of questions. When should the data for a given year be archived? The typical rule of thumb is that the three most recent years of data should be maintained “on-line.” It may be desirable, however, to maintain a longer time series if there is a keen interest in performing periodic follow-ups on the performance of past cohorts. There will also have to be a decision on how long the data will be kept in the archive. Will the data be archived for a specific period of time and then destroyed or will it be maintained indefinitely? Should the data be archived at the data warehouse or at the individual agencies? If the information is archived at the data warehouse, it will be readily accessible as needed to perform analyses of long-term trends. Agencies, however, may feel, out of a sense of ownership, more comfortable archiving their own data.

Data Transmittal, Storage and Archiving procedures
What format should the data be transmitted in?
What database structure should the data be stored in?
How long should the data be archived?
Should the data be archived at the data warehouse or at the individual agencies?

The Sequence for Addressing Issues

While all of these issues are important to the establishment of a data warehouse, two of the issues need to be addressed as soon in the process as possible. These are: (1) whether there is the need to have program participants sign informed consent agreements, and (2) the ownership of the merged database. If the data warehouse is established without informed consent agreements and it is later determined that such agreements are required to share participant data, the merged database will have to be destroyed. Further data collection will have to be postponed until informed consent agreements can be drafted, distributed to the field, and signed by program participants. It is also important to quickly resolve the issue of the “ownership” of the merged database. The question of ownership has important implications for the decisions regarding the access and use of the information. The sooner the issue of ownership is resolved, the sooner these other issues can be addressed.

Conclusion

The Blueprint has discussed the key steps and in creating integrated performance information for state level policy makers. Having reviewed the Blueprint, states may be considering how to proceed given federal initiatives in this area, including OMB's common measures, DOL's EMILE System, and the pending reauthorization of WIA, Perkins, and other federal acts related to workforce development.

States may choose to proceed in a number of different ways. The development of a shared information system is something that should be useful no matter the outcomes of the federal initiatives. Electronically linking records from multiple programs with files containing outcome data will facilitate the implementation of common measures and reporting the performance information likely to be required by the reauthorized acts, as well as useful in meeting state-identified needs.

If states find some or all of the performance measures recommended here to be useful, states can implement them as additional measures to those necessary to satisfy federal requirements. Experience has shown, moreover, that federal performance measurement requirements evolve over time. If states find the IPI measures to be useful, the measures may be reflected in future generations of federal acts and guidelines. Federal performance requirements have a powerful effect on program implementation and results; if states find the IPI measures to be useful, it would be very helpful if future federal requirements were aligned with the measures.

Finally, no matter the specific course of events in Congress or the federal agencies, the basic issues of building the capacity for and a culture of shared accountability are likely to be challenges that remain with states. Hopefully the state experiences and lessons learned that have been shared here, will help states as they continue to face these challenges.