

Putting Data to Work: Interim Recommendations From The Benchmarking Project

by Marty Miles, Sheila Maguire, Stacy Woodruff-Bolte and Carol Clymer

Performance Management: Hard Realities on the Ground

For Alice, a staff member at Opportunities for a Better Tomorrow (OBT), data entry is a long and tedious process. Data entry usually is. But the programs offered by this workforce development organization in Brooklyn are funded through five contracts with various government agencies, each requiring OBT to report information using a different online database. For each contract, Alice or one of her colleagues enters data on program participants' backgrounds and demographics, the services they receive and their job placement or educational outcomes. Each contract has crucial reporting deadlines to meet, but waiting for the online systems to "refresh" after each data field is updated sometimes makes it a slow endeavor.

Alice is careful with her data entry, because in a single OBT program one participant's services might be funded through a government contract while services for another are paid for by a foundation grant. OBT staff also have to make sure that backup paper documentation is in participant files, ready for possible contract audits. As time-consuming as this process may be, Alice and her colleagues understand its importance. The funds from these public contracts—supplemented by private

The Challenge

Given the fragmentation of the workforce development system, it is difficult for funders, policymakers and practitioners to know what outcomes constitute "good" performance. Furthermore, the diverse reporting requirements of workforce development funders take significant time and energy to navigate, thus sapping frontline providers' capacity to use data for program improvement.

Key Recommendations

Policymakers and funders at every level—but particularly the federal level—need to better support a culture of continuous learning and improvement across the entire system. To do this, they should:

1. Move toward more consistent definitions of performance measures.
2. Implement new technology or adapt existing systems to allow funder and program databases to exchange information more easily.
3. Provide more useful reports for practitioners about local and state data trends.
4. Offer more opportunities for program providers to learn from existing research and from their peers.
5. Encourage more programs to participate in The Benchmarking Project to enhance the field's ability to define "good" performance and to strengthen performance improvement efforts across the system.

foundation grants and other donations—are what make it possible for OBT to provide general equivalency diploma (GED) preparation, literacy services, business skills training or job placement assistance for 2,000 area residents each year. But the inefficiency of entering data into multiple systems—as well as OBT’s own internal database—is frustrating.

OBT’s extensive reporting processes reflect the national drive over the past 30 years toward greater accountability for programs receiving government and philanthropic funding. Modeled on the private sector’s emphasis on performance standards, grants typically require tangible, measurable results—and outcomes data play a major role in decisions about continued support. Contract payments tied to specific deliverables, such as job placement and retention, are often essential to organizations’ cash flow and bottom lines.

While organizations face external pressure to demonstrate results with outcomes data, most also realize that it is critical to use data to strengthen program services and spend resources effectively. As described in Public/Private Ventures’ 2006 report *Good Stories Aren’t Enough: Becoming Outcomes-Driven in Workforce Development*, frontline practitioners are increasingly engaged in a “cycle of continuous improvement,” analyzing individual data to glean knowledge that can lead to better results. This process is driven by the following questions:

- **What are our results?** Are we serving our target population? How many participants are showing improved skills? How many are securing employment and sustaining that employment for the long term? Are we seeing an increase in participants’ income and benefits?
- **What is “good” performance?** How do our results compare with others? How are our results changing over time? What level of performance should we expect?

The Benchmarking Project

- The Benchmarking Project began in 2004 with intensive work in three cities to understand the types of data local programs were collecting and related performance management issues.
- We designed and piloted a web-based survey to capture aggregate data from programs about participant demographics, services and outcomes for a recent one-year cohort of enrollees.
- As of March 2010, 214 programs from 159 organizations had submitted data and additional programs are joining all the time.
- Organizations receive confidential reports that allow them to compare their programs’ job placement and retention results (anonymously) with programs that share similar characteristics (e.g., size of cohort or type of service strategy). In each case, the median outcomes serve as initial benchmarks of performance.
- Organizations voluntarily decide to participate in The Benchmarking Project, and programs can update job retention data or submit new surveys on other cohorts at any time. There is no cost to organizations except the time needed to respond to the survey. Participating organizations indicate that the project’s “apples to apples” reports are a useful tool for identifying areas of program strength as well as areas needing improvement.
- To support program improvement efforts of participating organizations, The Benchmarking Project provides workshops, webinars, online discussions and other resources related to performance management and effective practice.

- **How can we improve?** How do results vary by type of service or participant? What can we learn from peers and research? What do answers to these questions tell us about potential changes in our strategies?

Similar questions about performance exist for workforce development funders and policymakers. The desire to answer them—particularly the “what is ‘good’ performance?” question—inspired the Annie E. Casey Foundation to launch The Benchmarking Project in partnership with Public/Private Ventures (P/PV). The project’s long-term goal is to identify realistic performance standards for programs in the workforce development field. The project is now completing its pilot phase after three years of data collection, and as of March 2010, 214 programs from 159 organizations across the country had submitted confidential data about their programs.

The Benchmarking Project has already provided compelling information about how specific program characteristics correlate with different outcomes and what some reasonable “benchmarks” might be.¹ While this brief includes initial benchmarks gleaned from the 200-plus programs in the current sample, these findings cannot yet be generalized—because the sample is not yet “representative” of the field. The next phase of the project seeks to engage a larger sample of organizations that better represents the many different types of programs and service providers operating across the country, with enhanced technology to accommodate more participation.

Yet even as The Benchmarking Project continues to expand, it already offers important insights about what service providers “on the ground” experience as they seek to better understand and improve their outcomes. This brief highlights issues that currently make it difficult for providers to fully engage in a cycle of continuous improvement and shows how the Benchmarking approach can help advance the use of outcomes data in the workforce field. We hope this information will be useful for policymakers and funders as they consider ways to more effectively support the improvement of workforce development programs—and ultimately produce better results for the participants and businesses they serve.

The Issues

“What Are Our Results?”

According to the National Conference of State Legislatures, at least 12 different federal agencies provide funds for workforce development programs—including the departments of Labor, Health and Human Services, Education, Housing and Urban Development, Agriculture, Transportation, Commerce, Energy and Veterans Affairs—and those funds are then channeled through state and local agencies.²

The programs in The Benchmarking Project’s national sample reflect this complexity. As of March 2010:

- 62 percent of participating programs had at least two different sources of funding.
- 25 percent had US Department of Labor funding through the Workforce Investment Act (WIA).
- 27 percent were funded through sources related to the Temporary Assistance for Needy Families (TANF) legislation administered by the US Department of Health and Human Services.
- 18 percent were funded through state vocational rehabilitation agencies related to the US Department of Labor or US Department of Veterans Affairs.
- 35 percent were funded through other federal, state or local government initiatives.
- 47 percent had support from foundations or other private sources, such as local United Ways.

Organizations like OBT and others participating in The Benchmarking Project are the “connect points” where federal, state or local dollars finally land to address workforce development needs. Many organizations piece together funds from multiple public sources to help targeted populations build skills that can lead to quality jobs. Some seek additional help from private foundations or United Ways to close remaining budget gaps. Each government contract or private grant can bring a different set of assessment processes, expected outcomes and data reporting requirements. These differences make it hard for funders and practitioners alike to understand the overall results of an organization’s program that is funded by multiple sources. Program leaders also find that their staffs are stretched by the time demands of entering data into multiple databases—a situation that owes to differences in how expected results are defined as well as the difficulty of sharing data between diverse technology systems.

Issue #1: Inconsistent definitions of outcomes across funding streams make it difficult for decision-makers and practitioners to get a clear picture of overall program results.

Leaders in organizations participating in The Benchmarking Project work to create an outcomes-focused environment in which staff have a clear understanding of current program results. But this effort is often complicated by differences in how outcomes are defined, including differences in what is being measured, when it is measured and who it is measured for. The Benchmarking Project’s

survey asked programs if they collect data on a number of outcomes (e.g., retention at 3, 6 and 12 months) and also asked a variety of questions to clarify how enrollment, placement and retention results are defined. The results illustrate how definitions are often shaped in response to specific funder guidelines. Differences in *what is being measured*, for example, can be seen in the myriad ways programs defined a “job placement”:

- 64 percent of programs counted temporary employment as a valid placement; 36 percent did not.
- 40 percent required that participants work a minimum number of hours in jobs for them to count toward placement goals; 60 percent of programs had no minimum requirement.
- 43 percent required that participants earn wages above a specific threshold to be counted as placed; 57 percent did not.

Definitions also differed for “job retention”:

- Close to one quarter of programs used the “snapshot method” when measuring three- or six-month retention—e.g., the participant was verified to be working on or after the 90th day after placement (or for WIA-funded programs, in the third quarter after exit from the program).
- Roughly one half of programs counted someone as “retained” if he or she worked continuously or almost continuously for the retention period, whether or not it was for the same employer.³
- One fourth of programs required a participant to work continuously for the same employer to be “counted” as retained.⁴

The Benchmarking Project sample also shows differences in *when participants are counted* and *when results are measured*. In 50 percent of programs, a person was considered “enrolled” on the first day he or she attended a program activity. Meanwhile, 36 percent of programs defined participants as enrolled only after they completed five or more days of activity. Likewise, while many programs counted someone as “placed” after one day of employment, some programs receiving TANF

Data Analysis Process for The Benchmarking Project: Comparing “Apples to Apples”

One of The Benchmarking Project’s goals is to create a process that will allow programs to compare key outcomes, such as job placement and retention, to those of other programs with similar characteristics. We used a statistical process to create meaningful peer groups:⁵

- We considered program population, services and other characteristics that might influence program outcomes (e.g., the size of the program cohort enrolled in a particular year).
- For each characteristic (e.g., size of the cohort), we tested various subgroups of programs (e.g., those with smaller versus larger cohorts) to see if there were statistically significant differences in outcomes between the subgroups.
- If significant differences in outcomes were found, we compared programs in one subgroup or category only with programs in that same category, since those in another category would not be expected to perform at the same level.
- Organizations participating in The Benchmarking Project then received confidential reports showing their placement and retention outcomes compared with the median rates for their particular subgroups.

funding through state agencies required participants to be on the job for 30 to 90 days before programs were credited for accomplishing a placement outcome.

Finally, differences in *who is being counted* also make it difficult to get a consistent picture of results. At a general level, as programs communicate their performance “rates,” knowing who is included in their calculation is critical. Some programs, when computing job placement or retention percentages, only include participants who have completed planned activities. Others have decided that it is important to count all enrolled participants in the calculation, including drop-outs. In some cases, the decision about which participants to include is driven by funder requirements.

Why do these inconsistencies in definition matter? When evaluating or comparing program outcomes, it is often hard to separate meaningful differences (i.e., related to program effectiveness) from differences that are tied to how outcomes are defined. As shown in Table 1, The Benchmarking Project’s current findings indicate that programs’ ability to select their participants has a statistically significant effect on some outcomes: Programs required to accept all eligible applicants had much lower median placement and six-month retention rates than programs that selected participants based on criteria they established themselves. Programs that were only “partially” selective—because of differing funder requirements or a decision to be less selective to meet enrollment goals—had median rates between those of the other two groups.

Table 1: Programs’ Ability to Select Participants: Effect on Outcomes

Degree of Program Selectivity	Median Job Placement Rate	Median Six-Month Job Retention Rate (Out of Number Placed)
Full Selectivity (n=77)	64%	62%
Partial Selectivity (n=104)	46%	59%
No Selectivity (n=33)	37%	48%

Note: Data drawn from March 2010 analysis of 214 programs in The Benchmarking Project. “Full selectivity” means that programs had the ability to select participants based on certain criteria, and they were selective. “No selectivity” means that programs were required to accept all individuals who showed up for services. “Partial selectivity” means either that programs had the ability to be selective but in fact had accepted all or most applicants, or that because of multiple funding source requirements they could be selective with some participants but not with others.

The Benchmarking Project is also analyzing data that explore how differences in definition might affect outcomes. For example, programs using a “snapshot” definition for six-month retention showed a significantly higher median retention rate (62 percent) than programs defining retention as continuous employment (56 percent). While further analysis will help answer other questions about the effects of definitional differences, it is already clear that the lack of consensus about how to define outcomes makes evaluating overall performance a complex endeavor.

Issue #2: The customized technology systems of different funders require duplicate, time-consuming data entry for practitioners, often without providing useful reports in return.

Funders (especially those in the public sector) have sought to streamline reporting by designing or purchasing web-based databases that grantees use to submit participant and outcome information. But the very systems designed to support an increased emphasis on performance often serve to distract programs from focusing on improving results. For participants in The Benchmarking Project, the two most frequently cited challenges associated with using data for program improvement are:

- The time it takes to enter data into multiple funder databases.
- The difficulty of retrieving information from these systems that could help with program management or improvement.

In OBT’s case, for example, each of its government databases has a few standard reports that can be printed, usually consisting of only what OBT staff have entered. Staff at OBT would like to know how their outcomes have differed over time for various subgroups (e.g., participants of a particular demographic or those receiving specific services). But they cannot query their funder databases for answers to these questions. Nor can OBT download the data into its internal database for further analysis on its own. As another program operator notes, the experience is like “feeding data monsters that gobble up information in a one-way communication process.” Other providers add that even retrieving summaries of data they have submitted sometimes requires working through additional bureaucratic steps.

Thus, most programs must maintain separate internal spreadsheets and databases to have the flexibility to use their data for program improvement. However, these in-house systems cannot “talk” with funder databases to make data sharing easier—and when programs must choose where to spend time, keeping funders’ systems up-to-date often takes first priority because these are linked to contract requirements—as well as payments.

Similar reporting issues exist for programs with multiple private funding sources. Providers document foundation-grant results through a variety of reporting forms, external databases and third-party evaluations. As one provider notes, the “cost ratio” of the time required for reporting to private funders relative to the size of those grants is often much higher than for government contracts.

“What Is ‘Good’ Performance?”

As funders and policymakers review outcomes data to evaluate program success and make funding decisions, how do they judge programs’ results? Programs may be able to document improved outcomes over time or that they have met annual contract goals. But finding external points of comparison in the field that are meaningful and relevant—as is done when private corporations set benchmarks for their performance with peer organizations—is a bigger challenge.

Issue #3: Performance standards and comparisons that do not take program differences into account make it hard to know what constitutes “good” performance.

One of the most striking realities about the workforce development programs participating in The Benchmarking Project is their diversity.⁶ In terms of program type and design:

- Some served small numbers of individuals in a year’s time in an intimate, supportive environment; others served several thousand in a year.
- 44 percent of programs in the sample offered some sort of occupational skills training to a majority of their participants, while others were designed to offer primarily job readiness training and job search assistance.
- Program settings varied; about half the sample’s programs were housed within organizations that specialized in workforce development, while the other half were part of multiservice agencies.
- Programs were spread across the country, with employment and wages affected by different regional economies.

In terms of target populations, some programs appeared to be “generalists,” serving a variety of hard-to-employ individuals, while others seemed to focus on one or two populations:

- The majority of participants at 22 percent of programs were formerly incarcerated.
- The majority of participants at 20 percent of programs reported having a disability.
- The majority at 16 percent of programs were TANF recipients.
- The majority at 10 percent of programs were young adults ages 18–24.

Participants' education and literacy levels also varied across programs:

- In 18 percent of programs reporting participant education levels, at least half of the participants lacked a high school diploma or GED.⁷
- Of programs reporting assessed reading levels, 9 percent served cohorts in which at least half of the participants were reading at or below the sixth grade level, while in 21 percent of programs at least half were reading at or above the tenth grade level.

In spite of this diversity, funders and policymakers frequently assess the performance of funded programs by looking at them as one group. They may communicate results through “report cards” that spotlight program outcomes but do not provide additional data related to differences in populations served, service strategies used or other significant program characteristics.

The Benchmarking Project's data suggest that such differences have a significant effect on outcomes. The size of a program cohort, for example, correlated to statistically significant differences in placement and retention rates. As seen in Table 2, smaller programs—those enrolling 100 or fewer participants during a one-year period—reported higher median placement rates as well as three- and six-month retention rates than those enrolling more than 100 participants in a year.

In some cases, the population served also had a statistically significant effect on outcomes. For example, programs in which more than 50 percent of participants were receiving TANF had a lower median job placement rate than programs with fewer TANF recipients.

Finally, the types of services offered—and the percentage of participants receiving those services—made a difference in the outcomes achieved. Programs providing vocational skills training focused on an industry-recognized certification or credential to more than 50 percent of participants reported significantly higher placement and three-month retention rates than programs providing certification training to fewer participants. Internships were another program service that yielded differences in job retention outcomes: Programs that connected 75 percent or more of their participants to internships had significantly higher median retention rates at both the three- and six-month points.

Table 2: Program and Participant Characteristics Affecting Performance

Characteristic	Median Placement Rate	Median Three-Month Retention Rate (Out of Number Placed)	Median Six-Month Retention Rate (Out of Number Placed)
Program Size			
1–100 participants served	65%***	80%***	63%*
More than 100 participants served	45%***	71%***	56%*
Receipt of Temporary Assistance for Needy Families (TANF)			
50% or fewer receive TANF	55%***	75%	61%
More than 50% receive TANF	42%***	74%	52%
Vocational Skills Training Leading to Certification			
50% or fewer enrollees participate in skills training leading to certification	48%***	72%***	56%
More than 50% participate in skills training leading to certification	65%***	83%***	67%
Internships			
75% or fewer enrollees participate in an internship	51%	74%**	56%*
More than 75% participate in an internship	62%	86%**	74%*

Data drawn from March 2010 analysis of 214 programs in The Benchmarking Project.

* p<0.10; **p<0.05; ***p<0.01

The reports received by Benchmarking Project programs have allowed them to compare their outcomes to programs that share these characteristics and others (e.g., smaller programs compare their results to other smaller programs), and programs have indicated that the comparisons offer valuable information about areas for improvement. Benchmarking Project data already indicate that expectations for what constitutes “good” performance must take into account key program differences. Without utilizing this kind of information, funders can set high across-the-board goals that may unintentionally drive program providers to recruit participants who are “less in need” than those normally targeted as part of their organizational mission.

Under the federal Job Training Partnership Act (JTPA) in the 1980s, the outcomes of local workforce programs were adjusted through a system-wide analysis model that took into account varying economic and population factors. However, while WIA has implemented a performance negotiation process that factors

in the percentage of low-income individuals in an area, it has not adequately adjusted its performance standards to reflect the characteristics of the population served.⁸ Without data that put results in the context of relevant comparison groups, it is difficult to accurately gauge the effectiveness of programs that initially appear to be high or low performing.

Within The Benchmarking Project sample, programs that served out-of-school and out-of-work young adults ages 18–24 voiced this frustration regarding contextual data loudly. This population’s progress can be slow and zigzagged because of the wide range of literacy, family and life skills issues that often need to be addressed. Funders assessing GED attainment or employment rates for programs serving young adults more broadly may make “apples to oranges” comparisons if they do not account for differences in the subgroups served by these programs (e.g., youth with very low literacy levels versus youth with higher levels). It can take many months or even years for a low-skilled young adult to arrive at a point where achieving a GED, securing a job or maintaining steady employment is a realistic expectation. But funders and practitioners often do not have an agreed-upon set of interim milestones to use as standards for evaluating shorter-term progress.

“How Can We Improve?”

As organizations participating in The Benchmarking Project get a clearer sense of their results—and how “good” these results are compared with similar programs—they have begun to focus on ways to achieve better outcomes. They understand that performance improvement requires a change in their organizations’ “data culture.” But their experiences also suggest that funder expectations of accountability need to be better balanced with support for program leaders as they work to make these changes.

Issue #4: Practitioners need increased support to develop and maintain an internal culture of continuous improvement.

Even though data submitted by programs in The Benchmarking Project are kept confidential and comparisons are anonymous, programs may view submitting their data as a risk, since they are unsure how they will fare in national comparisons. Yet more than 200 programs have been willing to take this risk, because of a strong incentive to glean ideas from the data and from their peers about how to improve performance. This incentive derives both from their strong sense of “mission” and from the significant “bottom line” pressure to meet the goals of outcomes-based funding contracts.

In workshops and forums sponsored by The Benchmarking Project, program managers have shared the challenges they face in strengthening their internal data culture. They must perform a delicate balancing act: to simultaneously require that staff use data to demonstrate results while creating a nonjudgmental learning

environment that views data as a tool to help staff provide better service. Managers must also do this at a time when program staff are stretched to meet the growing needs of participants and have little time to devote to working with data.

This management balancing act is not easy, and program leaders welcome opportunities to hear about effective data practices from their colleagues in the field. They ask for suggestions about more user-friendly database systems that can be adapted to their varied reporting needs. They seek ways to increase communication among staff in various roles about factors potentially influencing results. Programs also want help in developing strategies for managing performance and helping frontline staff make the connection between individual results and overall program goals. They ask for increased access to research about effective program strategies and practical models to inform their thinking. Finally, program leaders are looking for financial or in-kind resources to support updating databases and training staff in their use as well as conducting more in-depth data analysis.

Recommendations

Practitioners taking part in The Benchmarking Project are working hard to implement cultures of continuous improvement within their organizations. But these organizations are only one set of players in a multitiered workforce development system that includes many federal, state and local funders. For overall outcomes to improve, a similar kind of culture has to evolve system-wide. Just as program managers must balance accountability for results with support for learning and improvement across their organizations, public and private funders need to encourage that balance within the larger system. Funders and policymakers at every level can play a critical role as catalysts of a new kind of partnership—by developing better ways to share information and working together with practitioners to improve the results of the workforce system. Adopting this spirit of increased collaboration could draw attention to the issues identified in the previous sections and illuminate possible solutions. This kind of culture shift will take time—but funders and policymakers should start by taking action in the following areas:

1. **Move toward more consistent definitions of performance measures.**

The workforce field needs more consistent definitions of key performance measures across its many funding sources. Not only would this simplify the reporting process for individual programs, but it would also enhance the field's ability to gain access to and analyze comparable data to support broader learning. Groups of private foundations in cities like Baltimore and New York are currently seeking to better align their definitions of performance outcomes for local workforce development grantees, an important first step that could lead to the sharing of definitions across other networks of workforce funders. But to effect real change, leadership also needs to come from the public sector—ultimately at the congressional level, given that federal legislation drives most state and local public measures.

Further research is also needed—to identify appropriate interim milestones for programs that work with populations requiring long periods of time to achieve workforce development goals (e.g., young adults with low literacy skills). This information will allow practitioners and funders alike to better evaluate program effectiveness in the short term, while providing more “real time” measures to guide program improvement efforts.

2. Implement new technology or adapt existing systems to allow funder and program databases to exchange information more easily.

While technology by itself will not sustain a culture of continuous improvement, it is a critical element in such a process. The many systems used across state and local entities to track workforce program performance must be better able to “talk” with each other and with the databases of local providers. At a minimum, local or state public funders need to facilitate program providers’ ability to export data from their internal databases into funder systems and to retrieve useful summary data from those systems. This capability would reduce time-consuming data entry and could free up staff to focus on distilling learnings that could improve practice. It could also enhance data sharing across various parts of the workforce system to reveal trends and inform research on performance benchmarks.

3. Provide more useful reports for practitioners about local and state data trends.

It is not unusual for local or state workforce funders to produce report cards showing grantee results or rankings on a number of factors. To some extent, the sense of “competition” these reports engender can provide helpful motivation to perform. But to improve outcomes, frontline providers need access to additional data that can help them understand the factors contributing to their results. State and local funders could assist program operators by providing more cumulative reports that summarize trends in outcomes related to participant demographics, services provided, targeted industries and occupations, and so forth. They could also make their data analysis staff or other resources available to programs that seek answers to more specific questions.

4. Offer more opportunities for program providers to learn from existing research and from their peers.

Public and private funders at federal, state and local levels need to support a collaborative learning environment by engaging grantees in regular discussions about participant and service trends and the factors grantees think are affecting performance. They could sponsor more webinars or other events that highlight existing research and feature programs using effective practices. Funders could also implement capacity-building models like the workshops and forums used in The Benchmarking Project to help management and frontline staff strengthen their ability to use data as a performance improvement tool.

5. Encourage more programs to participate in The Benchmarking Project to enhance the field’s ability to define “good” performance and to strengthen performance improvement efforts across the system.

As the workforce development field evaluates outcomes to identify effective programs and set performance goals, it needs a better way to incorporate evidence about the program characteristics that affect results. The approach used by The Benchmarking Project offers the field a feasible model for creating more “apples to apples” comparisons across a diverse set of providers and for identifying evidence-based performance standards. But for this tool to provide even more useful information about strategies and benchmarks for funders, policymakers and practitioners, a much larger number of programs will need to be engaged.

The Benchmarking Project: Next Steps

In the project’s next phase, a larger data sample will enable us to make more refined comparisons and—in the future—will allow us to more accurately define performance benchmarks for programs with various combinations of characteristics. We are also working to understand the total scope of providers in the workforce development field to get a better sense of the number and array of programs needed to create a representative sample for further research purposes. Finally, we will gather additional data from high-performing programs in the sample to identify specific strategies or practices that merit more in-depth case studies or rigorous evaluation.

As we move forward, we will also deepen our efforts to connect participating organizations through a “learning community” designed to support their continuous improvement efforts. With funding from local foundations, more than 50 organizations in New York City and Chicago are involved in regular group forums and technical assistance to help program leaders establish a stronger staff “culture” around using data. As part of these efforts, we will work with local funders to explore using more consistent definitions of performance measures and will work with providers to assess the correlation of various interim milestones with longer-term outcomes. We hope to do similar work with practitioners and funders in other cities.

For the national learning community, we plan to conduct frequent webinars that will explore trends emerging from the project’s data as well as how various organizations are using data as a performance management tool. The Benchmarking Project’s enhanced website will link organizations to research studies and other resources related to effective workforce program strategies. Discussion forums and online conferences will focus on critical issues around data and practice identified by participants.

Conclusion

Our experience in the first phase of The Benchmarking Project has shown that many service providers, such as OBT, are trying to navigate complex and fragmented funding streams so that they can provide services—and trying to answer the basic question: “What are our results?” Simplifying the process through more consistent definitions and better technology is critical. But it is equally (if not more) important that the workforce system of funders and providers focus on the two other questions related to improving results: What is “good” performance, and how can we improve?

We believe The Benchmarking Project’s approach to assessing relative program results offers a cost-effective way to help funders, policymakers and practitioners answer the “what is ‘good’ performance?” question. The more programs that participate in the project, the more compelling and generalizable the data will become. However, helping programs focus on “how can we improve?” is even more challenging. An unintended consequence of the complexity of workforce development funding is that it saps the resources and capacity of the organizations on the front lines, hindering their efforts to do the hard work of answering this final question. It is critical that policymakers and funders interested in improving overall workforce development results support better sharing of data trends among the system’s many players and offer more collaborative learning opportunities that tap the valuable experience and perspective each partner has to offer.



THE
BENCHMARKING
PROJECT

PUTTING DATA TO WORK

More information about
The Benchmarking Project is
available at www.ppv.org or by
emailing benchmarking@ppv.org.

Endnotes

1. The Benchmarking Project is using the “median” performance rate (midpoint of the data range) as an initial benchmark.
2. National Conference of State Legislatures. “Federal Workforce Development Programs.” Retrieved 9/21/2010 from <http://www.ncsl.org/Default.aspx?TabId=13343>.
3. Forty-seven percent of programs counted participants as retained if they worked continuously or almost continuously for the three-month retention period, and 53 percent used this definition for six-month retention.
4. Twenty-eight percent of programs counted participants as retained if they worked continuously for the same employer for the three-month retention period, and 23 percent used this definition for six-month retention.
5. Analysis of variance (ANOVA), a statistical process for comparing group mean values, was used to determine whether statistically significant differences between categories existed.
6. To be eligible for participation in The Benchmarking Project organizations must directly provide workforce development services and track job placement and retention outcomes. They also must: have served individuals 18 or older, not including incumbent workers; have enrolled a cohort of at least 25 participants (not counting carry-ins) during a recent one-year period; provide complete aggregate outcome data for that cohort on job placement and retention (retention data at 6 and 12 months after placement is preferred; 3-month retention data will be accepted if longer-term data is not available); and provide summary information on cohort demographics, type and length of program services, and program budget.
7. Twenty percent of the programs participating in The Benchmarking Project were unable to report participants’ highest level of education.
8. Center for Law and Social Policy. 2009. “Recommendations for Reauthorization of the Workforce Investment Act Adult Program.” Center for Law and Social Policy: Washington, DC.



Public/Private Ventures

Philadelphia Office

2000 Market Street, Suite 550
Philadelphia, PA 19103

Tel: (215) 557-4400

Fax: (215) 557-4469

New York Office

The Chanin Building
122 East 42nd Street, 42nd Floor
New York, NY 10168

Tel: (212) 822-2400

Fax: (212) 949-0439

California Office

Lake Merritt Plaza, Suite 1550
1999 Harrison Street
Oakland, CA 94612

Tel: (510) 273-4600

Fax: (510) 273-4619

www.ppv.org

November 2010