NRS Data Monitoring for Program Improvement

By:

Larry Condelli
Peggy Seufert
Steve Coleman

AMERICAN INSTITUTES FOR RESEARCH
1000 Thomas Jefferson Street, N.W.
Washington, DC 20007

This guide was prepared for the project:

Promoting the Quality and Use of National Reporting System (NRS) Data
Contract # ED-01-CO-0026

U.S. DEPARTMENT OF EDUCATION

Office of Vocational and Adult Education, Division of Adult Education and Literacy

Susan Sclafani, Assistant Secretary for Vocational and Adult Education

Cheryl Keenan, Director
Division of Adult Education and Literacy

Mike Dean, Program Specialist
Division of Adult Education and Literacy

April 2004

Contents

Page

Chapter 1. NRS Data and Program Performance 1
Data-Driven Program Improvement 1
A Model for Performance Standards, Monitoring, and Improvement 2
Data and Monitoring: An Example 4
Overview of the Guide 5

Chapter 2. Setting Performance Standards for Program Quality 7
Accountability Systems and Program Quality 7
The Central Role of Performance Standards 8
Models for Setting Performance Standards 9
Continuous Improvement Model 9
Relative Ranking Model 11
External Criteria Model 12
Return on Investment Model 12
Adjusting Standards for Local Conditions 13
Student Characteristics 14
Local Program Elements 15
External Conditions 16
Making Standards Work: Shared Accountability and Avoiding Unintended Effects 17
Shared Accountability 17
Appropriate Rewards and Sanctions 19
Data Monitoring and Local Contact 20
Summary: Setting Local Performance Standards 21

Chapter 3. Under the Data: Performance Measures and Program Processes 25
Under the Data 25
Educational Gain 26
Follow-up Measures 29
Retention 31
Enrollment 33
Quality Indicators as Program Standards 35

Chapter 4. Planning and Implementing Program Monitoring 37
The Value of Monitoring 37
Approaches to Monitoring 38
Desk Reviews 39
Onsite Reviews 39
Ways To Collect Data for Monitoring 40
Program Self-Review 40
Document Reviews and Environmental Scans 41
Observations 42
Interviews 43

Contents (Continued)

Page

Monitoring With Indicators of Program Quality and Data 45
Seven Steps for Developing a Monitoring Plan 49

Chapter 5. A Model for Program Improvement 59
The Program Improvement Process 59
A Continuum of Program Improvement 59
A Model of the Program Improvement Process 61
Plan Your Strategy 61
Develop a Program Improvement Team 61
Prioritize Areas Needing Improvement 62
Gathering and Selecting Promising Strategies 64
Taking Action 65
Implementing Change 65
Evaluate the Results 66
Documenting the Process: Lessons Learned 67
Guiding Questions 69

Bibliography 71

Appendix 73
Indicators of Program Quality and Self-Review Instruments 73
Program Documentation 75

List of Tables

Page

Table 2–1 Performance Standard-Setting Models..... 10

Table 2–2 Program Performance for Setting Standards Under Two Models..... 12

Table 2–3 Summary of Local Performance Standard-Setting Process..... 22

Table 4–1 Advantages and Disadvantages of Desk Reviews..... 39

Table 4–2 Advantages and Disadvantages of Onsite Reviews 40

Table 4–3 Data Collection Strategies for Monitoring..... 41

Table 4–4 Monitoring With Indicators of Program Quality and Data..... 46

Table 4–5 Steps and Guidelines for Monitoring Local Programs..... 49

Table 5–1 Continuous Program Improvement Rubric 60

Table 5–2 Program Improvement Team 63

List of Exhibits

Page

Exhibit 1–1 Data-Driven Program Improvement....3

Exhibit 1–2 Educational Gains for ESL Levels and Performance Standards....4

Exhibit 3–1 NRS Data Pyramid: Educational Gain....26

Exhibit 3–2 NRS Data Pyramid: Follow-up Measures (GED, Employment, Postsecondary)....29

Exhibit 3–3 NRS Data Pyramid: Retention....32

Exhibit 3–4 NRS Data Pyramid: Enrollment....34

Exhibit 4–1 Overall Strategy for Program Evaluation and Continuous Program Improvement....57

Exhibit 5–1 The Program Improvement Process....62

Exhibit 5–2 Sample Data Report Educational Gain—ESOL Levels....69

Chapter 1. NRS Data and Program Performance

The National Reporting System (NRS) is now well established as the Federal accountability system for the adult education and literacy program. It includes measures of student outcomes, including literacy gains, improved English proficiency, attainment of postsecondary credential and student advancement to further education. These measures provide evidence of student achievements from attending adult education programs, and serve as indicators of program performance. Reporting these data meets your grant requirements and provides funders and other audiences with indicators of program success in helping students improve their literacy skills and achieve their goals. These data also can tell you the areas where your program is doing well, which students succeed in your program, and what types of classes and instructional arrangements are most effective.

Unfortunately, many adult educators are not accustomed to thinking about data in these program improvement terms. Although we all know that data have an administrative function—to give the state and the "Feds" the data they want—we are much less familiar with the ways in which data can be used to understand program performance. We know that good programs show good results, reflected in data in some good way, but we may also believe that good work is sometimes imperfectly reflected in data. In this guide, we explore the connection between data and program quality, not only to understand it, but also to learn how to use data to improve program performance.

Data-Driven Program Improvement

Accountability requirements often make it appear that there should be a direct connection between outcomes and program quality. However, the relationship between data and program quality is more complex than it may first seem. Furthermore, the relationship is dynamic, as data, in the form of performance standards and other numeric goals, not only *measure* program performance but can *change* it as well. In this guide, we will explore the dynamics of performance and accountability by:

- Demonstrating the interrelationship of data and program performance—how data affect program performance and how performance affects data;
- Exploring ways to monitor programs to enhance the connection between performance and data; and
- Discussing ways to identify and implement program improvement efforts.

Specifically, we will examine how to use NRS data to affect performance through the use of performance standards. We will describe models for setting performance standards and the different policy and program improvement goals you can implement with each model. We also explore the program processes and variables that influence performance on each

measure, ways to examine these processes through monitoring, and models for program improvement.

A Model for Performance Standards, Monitoring, and Improvement

One of the main goals of this guide is to demonstrate that NRS data can drive and inform program improvement efforts. Exhibit 1–1 illustrates a model for a data-driven program improvement process around which we have organized the guide. The process includes the setting of performance standards, which serve as the numeric targets for outcomes and other measures. The targets define a level of acceptable performance that when linked with rewards and sanctions, motivate program processes and outcomes.

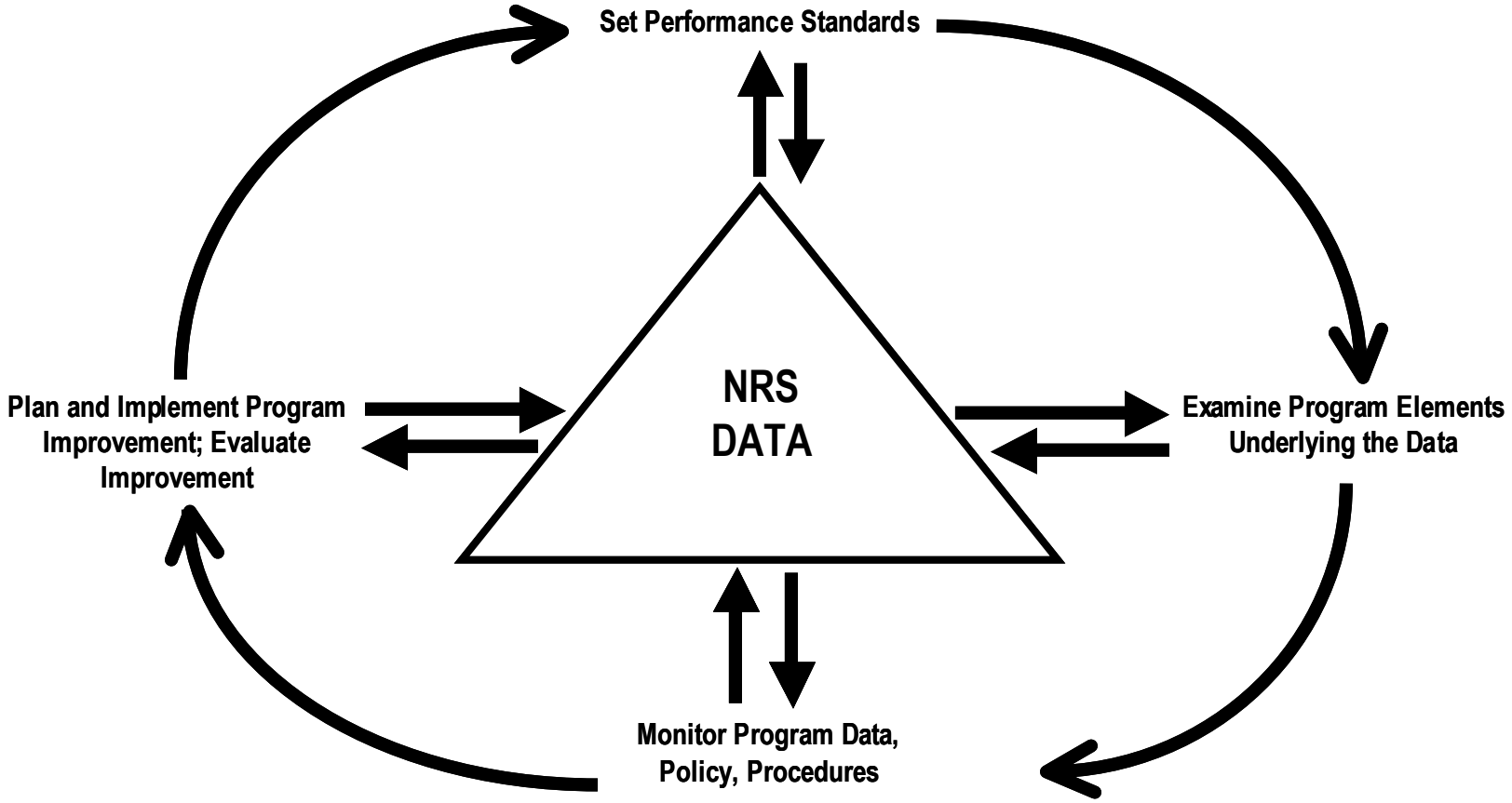
Underlying the performance measures are program elements that represent a series of policy, programmatic, instructional, and data collection activities performed by program staff. The measures represent the observable end result of these program processes. A program's performance on NRS measures tells us much more than how the program collects and reports the data. The data are also a mirror of what the program does—its intake, goal setting, instruction, and assessment methods, for example, are all reflected. NRS data are a gateway into the heart of the adult education program.

The data and their underlying elements guide program monitoring by identifying the areas of program operation to evaluate. However, because an adult education program consists of many different elements, it is not always clear from data alone which aspects of the program are effective and which need improvement. Through monitoring, both onsite at the program and through data and desk monitoring, you can more fully review program procedures and understand the program.

Through a review of performance data and program monitoring, you can identify areas of a program's strengths and weaknesses. This information will guide the planning and implementation of program improvement efforts. By focusing on the program activities that appear to affect performance, you can define specific actions that the program must take to inform performance. If successful, the desired outcome of these changed procedures is improvement in outcome measures and performance standards. Then, as the circular relationship in exhibit 1–1 illustrates, the improvement process can begin again.

The exhibit also shows that data are central to the program improvement process, informing all of the steps along the way. Data provide the measures for the performance standards, suggest an explanation for performance, and guide program-monitoring efforts. Once the program makes improvements, you can look to the data to assess whether the changes resulted in improvements in program performance measures. Changed performance will then affect the setting of future performance standards.

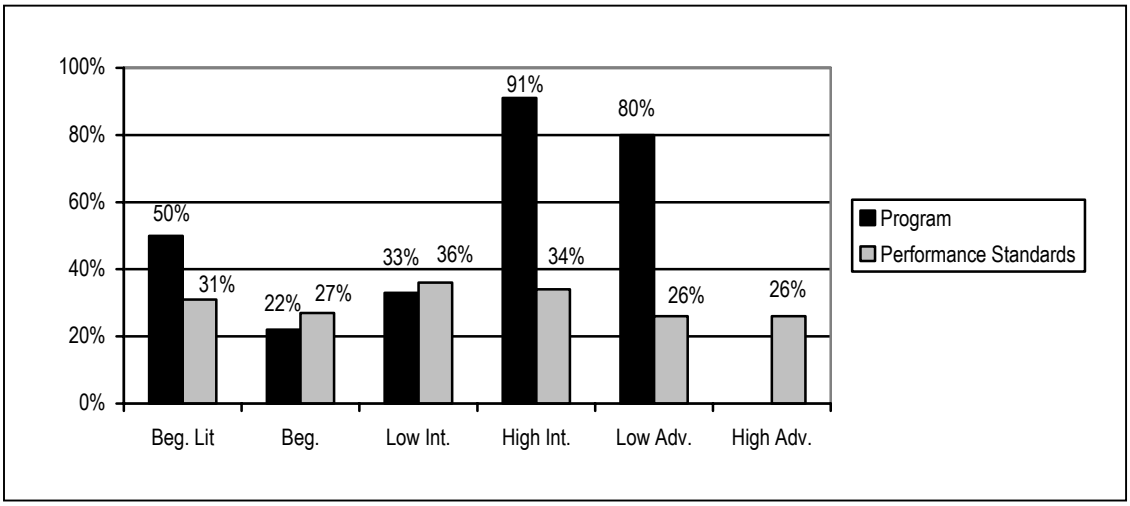
Exhibit 1–1
Data-Driven Program Improvement



Data and Monitoring: An Example

Exhibit 1–2 shows the percent of student advancements for each of the English as a second language (ESL) levels of the NRS for a local adult education program. The chart also includes the program's performance standards for each level. The program exceeded its targets for three of the ESL levels, failed to meet them for two levels, and could not meet its targets for the high-advanced ESL level because it did not serve any students at that level. By reviewing this chart you might conclude that the program needs to improve its performance on the levels that it failed to meet targets (beginning and intermediate levels).

Exhibit 1–2
Educational Gains for ESL Levels and Performance Standards



Although this conclusion is correct, our hope is that this guide will help you raise other questions when you see this type of chart, such as:

1. How were the performance standards set for this program? Were they based on past performance or some other criteria?

2. Are these standards appropriate given the pattern of performance the program has shown? For example, are the standards too low at the higher levels, where performance greatly exceeded targets?

3. Is this performance pattern similar to that observed in previous years? If not, what has caused it to change? Will this affect setting of performance standards in the future?

4. What are the program's assessment and placement procedures? What assessments are used for pre- and posttesting?

5. Is the program using the same assessment methods for high- and low-level ESL? If so, is this appropriate given the performance pattern?

6. What type of curriculum and instruction is the program offering? How does it differ by instructional level?

7. What are student retention patterns by level? Is retention affecting the differences in performance among students at different levels?

8. Could the program's recruitment practices have had an influence on performance? How many students is it serving at each level?

9. Why are no students enrolled at the highest ESL level? Is this a result of recruitment, type of classes the program offers, or placement procedures? Does the program need to change its recruitment practices?

In this guide, we will explore approaches and strategies that will enable you to organize and systematically address questions like these when you examine your local program data. The use of these strategies will help guide your program improvement efforts through the use of data and program monitoring.

Overview of the Guide

Adult education is facing a time of change. The Workforce Investment Act (WIA), the program's authorizing legislation, will soon expire, and new legislation is pending in the U.S. Congress. States may soon have to develop new state plans, are setting new state performance standards and face possible requirements for setting local performance standards and local program monitoring. Along with these changes are new opportunities for program improvement. The NRS has operated for 3 years, and states and local programs have improved their data collection processes and understanding of data. The U.S. Department of Education has supported state training and technical assistance on accountability throughout this time, focusing on NRS data quality and NRS data use. There are now 3 years of NRS data on which to base program decisions and understand performance.

This guide builds on these past efforts to promote program improvement through use of NRS performance data to guide program monitoring. It is organized into four additional chapters, each addressing one of the data-driven program improvement elements identified in exhibit 1–1. Chapter 2, "Setting Performance Standards for Program Quality," discusses the central role of performance standards as policy-setting tools in accountability systems. It describes four models for setting local performance standards, the policy strategy each model represents, and the advantages and disadvantages of each. It also discusses the need to adjust standards to local program conditions and to tie standards to rewards and sanctions. The chapter also discusses the problem of unintended consequences that sometimes result from standards and emphasizes the need for shared accountability.

Chapter 3, "Under the Data: Performance Measures and Program Processes," discusses the program processes underlying accountability data. Focusing on educational gains, follow-up measures, retention, and enrollment, the chapter highlights the key data collection procedures, program policies, and program elements that affect each of these measures. Program quality indicators are identified as a way to standardize program processes and tie them to data.

In Chapter 4, "Planning and Implementing Program Monitoring," we discuss strategies toward monitoring, highlighting the importance of local program buy-in, participation, and shared accountability. We compare desk monitoring and onsite monitoring, identifying the types of information that can be obtained from each procedure. The chapter presents an organized strategy for planning and conducting monitoring and describes different approaches to monitoring among states linked to indicators of program quality.

The guide concludes with Chapter 5, "A Model for Program Improvement," that describes an approach to making improvement in local programs. The model describes steps for planning change, reviews action steps programs must take to implement change, and discusses the need to evaluate the results of changes. The chapter stresses the importance of planning and implementing change and evaluating program improvement on a continuum. An appendix to the guide contains examples of state monitoring methods and a list of resources for further information about the topics and issues presented in the guide.

This guide is the third in a series of guides and resources developed through the project Promoting the Quality and Use of NRS Data with funding from the Office of Vocational and Adult Education (OVAE) of the U.S. Department of Education. The purpose of this project is to develop resources and provide training to improve the quality and use of data collected for the NRS. Separate training materials to accompany this guide will be available in the summer of 2004. This guide does not include a presentation of NRS policy, the elements of the NRS, or any of the issues surrounding the collection of NRS data. Readers interested in more information about these topics should consult the *NRS Implementation Guidelines*, *NRS Survey Guidelines,* and *Guide for Improving NRS Data Quality*, available from the U.S. Department of Education or online (www.nrsweb.org).

Chapter 2. Setting Performance Standards for Program Quality

Measuring student outcomes has become a fundamental part of all educational programs. Public schools, job-training, vocational, and adult education programs alike must meet performance targets set by state and local agencies for student learning gains and other outcomes. The main reason underlying these performance requirements is the belief that good student performance means high program quality: a good program is one that has good student outcome measures. However, the relationship between student outcomes and program quality is complex: student outcome data only *indirectly* measure the program and instructional processes that define program quality. Furthermore, skillful use of outcome measures also can affect program processes.

In this chapter, we explore the complex relationship between program outcomes and processes. To do this, we first review the art and science of program accountability systems, focusing on the central role of *performance standards* as a guiding force. We discuss how performance standards can both reflect and influence program quality. Next, we present models for setting performance standards and discuss the conditions under which performance standards most effectively affect program quality. We conclude with guidance and examples for using performance standards to promote program improvement.

Accountability Systems and Program Quality

The NRS, as in all systems of program accountability, provides measures that gauge whether a program is meeting its goals and mandates. The measures give funders and other audiences indicators of program quality that they use to judge the success of the program and whether it should receive support. An accountability system can measure quality accurately only when it contains the following four essential characteristics:

- An underlying set of *goals* that the program is to achieve;
- A common set of m*easures* that reflect the goals. The measures must be clearly defined and can be qualitative or quantitative;
- *Performance standards* tied to measures, which set a level or target of performance that programs must achieve; and
- *Sanctions or rewards* for programs, tied to performance.

The NRS embodies goals of the adult education program, such as literacy development, secondary credential attainment, continued learning, and employment. The core outcome measures of educational gain, passing of the general education diploma (GED) tests, entry into postsecondary education, and employment reflect these goals. Although states can add additional measures reflecting their own goals, these NRS core

measures are required of all programs and define the central features of the accountability system.

Although the NRS also includes performance standards and rewards and sanctions for performance, states have more latitude in this aspect of the accountability system. States must set their performance standards in collaboration with the U.S. Department of Education and may soon be required to set standards for their local programs. States are to monitor program performance on the standards and establish procedures for performance-based rewards and sanctions. The purpose of these requirements is to improve program quality. This performance standard process, when properly implemented, can indeed be a highly effective way to foster program improvement.

The Central Role of Performance Standards

A performance standard is a numeric target assigned to a measure that defines an acceptable level of performance. Usually the program funder ties a reward to the achievement of the performance standard and/or imposes a sanction for failure to achieve it. It is this performance standard–reward–sanction dynamic that gives accountability systems their ability to affect program performance and quality. The performance standard serves as a goal, and the reward or sanction is a motivator for the program in achieving that goal. The program makes changes that will affect its performance on the standard to reach the goal. Through this process, accountability systems not only serve as a way to measure program performance, but they can affect program processes and outcomes.

The performance standard-setting process also makes standards effective policy and program improvement tools. Setting standards appropriately is a way to implement policy and is also a way to measure whether these policies are effective. For example, emphasizing specific measures helps define what policymakers believe is important and focuses programmatic efforts on improving performance on these measures. The same measures then become the means by which program success is evaluated.

The clearest example of this process is the use of standardized test scores as a measure of student learning. By setting a performance standard on educational gain, with a reward or sanction tied to achievement of the standard, programs focus on producing high test scores. To achieve higher scores, programs make changes in their processes and procedures (e.g., recruitment, assessment, instruction). The state or funding agency then uses higher scores, or evidence of increased educational gain among students, as the criteria by which to judge program quality.

This view of how accountability systems affect program quality represents the ideal. In reality, we know things are rarely that simple. A good accountability system is a necessary but not sufficient way to improve program quality. Although the use of performance standards can result in programmatic changes, whether these changes will result in *improving* program quality depends on whether the standards are set appropriately, the types of rewards and sanctions offered, an understanding of what the

standards are intended to do, and the assistance provided to programs to help them meet the standards. The use of performance standards to affect quality is as much an art as it is a science, and is best viewed as a living process—one that needs continual care and attention to make it work.

Models for Setting Performance Standards

Your state has negotiated performance standards with the U.S. Department of Education for all of the core NRS measures. The standard setting was based on your state's past performance and reflects a model of continuous improvement. Your state also may have set standards for local program performance based on these statewide standards. Many states use the same state standards for local programs, setting the same numeric value for each measure. However, this type of standards-setting model is just one approach—and may not always be the most appropriate for improving program quality.

There are four main models for setting standards, each reflecting different policy goals and having a different potential impact on programs. To set standards, you should first adapt the most appropriate model for your state or program, and then if necessary, adjust the standards for local program conditions, such as the types of students who attend or the type of instruction offered. Once the standards are set, you must monitor and guide the effect of the standards on programs to ensure the standards affect quality in they way you intended. Performance standards have the potential for producing undesirable or unintended effects that adversely affect program quality. You also must manage these unintended consequences.

Table 2–1 summarizes the four performance standard-setting models: continuous improvement, relative ranking, external criteria, and return on investment. Each model reflects a different policy and a different approach toward improving program quality. The table describes the strategy each model implements, an example of its use, and its advantages and disadvantages. When setting local program standards, you may use different models for each program or use a combination of models.

Continuous Improvement Model

Probably the most popular model for setting performance standards is the continuous improvement model, whereby the standard is set according to a program's past performance. The program's performance over the last several years is reviewed, and a standard is set above the previous year's level or above its average. The level increase is determined by policy and what is realistic and desirable from the program. With this approach, every program's standards must be set individually, and all programs usually have different standards.

Table 2–1
Performance Standard-Setting Models

Model	Policy Strategy	Example	Advantage	Disadvantage
Continuous Improvement. Standard is based on program's past performance.	Designed to make all programs improve compared to themselves.	Program will show a 10 percent increase from last year in the number of students advancing to low–intermediate ESL.	Works well when there is stability and a history of performance on which to base the standard.	Ceiling is reached over time so that little additional improvement is possible.
Relative Ranking. Programs are ranked by score on the measure, and the standard is the mean or median performance.	Used for relatively stable measures, where median performance is acceptable.	Percent of students passing the GED in the program will be equal or greater than the state average.	Focuses program improvement mainly on low-performing programs.	Higher performing programs have little incentive to improve.
External Criteria. Standard is set by formula or external policy criteria.	Promotes adoption of a policy goal to achieve a uniform higher standard of performance.	Twenty percent of all students enrolled in ABE annually will enter community college.	Appropriate for strategies when large improvements are needed, often over the long term.	External goal set without considering past performance may be unrealistic and unachievable.
Return on Investment. Net value of the program is compared to cost.	Provides information on whether services are worth the monetary or resource investment.	A sufficient number of students from the program will enter employment to offset the cost of instruction (through taxes or welfare savings).	Best used for deciding whether a particular program or site is worth the resources spent.	May ignore other advantages offered by the program or site.

Many policymakers believe this strategy is the most effective approach for setting standards, since it requires all programs to improve continually, but each program's improvement efforts are judged only against its own past performance. Comparing the program against its own past performance controls for factors such as student characteristics, local conditions, and factors unrelated to program quality that might affect performance, assuming the program remains stable in these areas. Consequently, it is easy to evaluate whether increased performance on the standards is related to program improvements. For these same reasons, most program managers consider it a fair way to evaluate their improvement efforts.

To use this strategy effectively, however, the program must have data on its performance history to calculate the standard. When the NRS first began, most states and programs lacked accurate data on their past performance and had to guess the level at which to set their standards. In later years as NRS data became available, most states and the U.S. Department of Education adopted the continuous improvement model to set performance standards.

The major disadvantage of this model is that eventually, programs reach a ceiling where continued improvement is very difficult or no longer possible. There is ultimately

an upper limit for performance (at most, 100 percent, although usually much lower). Therefore, over time this strategy is no longer effective or possible. However, it can be an effective model when the accountability system is relatively new or when you can be confident that the program is not near its peak performance—in other words when you know there is room to improve.

Relative Ranking Model

The second most popular model for setting performance standards is relative ranking. In this model, programs are put in a rank order on the measure, and the median or mean becomes the standard. By definition, all programs have the same standard for the measure: half of the programs will fall below the standard and half will be above.

This model is effective when the goal is to maintain a stable, uniform, or very similar level of performance among all programs. Because performance is focused on the average, variation among programs on the measure lessens over time as program performance becomes increasingly similar. The disadvantage, however, for the higher performing programs—those already above the average—is that there is little incentive to improve performance. Because performance has already exceeded the standard, the staff may perceive that there is little advantage to further program improvement efforts. Unlike the continuous improvement model, where there is always pressure on the programs to improve, the relative ranking model challenges only half of the programs to improve.

Given the characteristics of the relative ranking model, it is most appropriate when you want to direct program improvement efforts to only the half of the programs that are below the standard. This approach may be desirable to create a more homogenous system or when resources for program improvement are limited and thus are directed only to programs most in need. The model also may work when you believe the higher performing programs are already at or near their peak or do not need further improvement.

Table 2–2 illustrates differences between the continuous improvement and relative ranking performance standard-setting models. The table shows one program's performance in advancing students from low-intermediate adult basic education (ABE) over the last 3 years and the state average on this measure. To set a performance standard on this measure for the program under the continuous improvement model, you would consider the program's latest performance, 42 percent. Because the program has improved 8 percent in 2 consecutive years on this measure, you might set it around 50 percent under this model. Under the relative ranking model, you would set the standard around the average of 38 percent. Under this latter model, the program already exceeds this standard and would only need to maintain its current level of performance to meet the standard next year. In contrast, under the continuous improvement model, the program would need to continue to improve.

Table 2–2
Program Performance for Setting Standards Under Two Models

Percent of Students Advancing From Low-Intermediate ABE			
	Program Year 2000	Program Year 2001	Program Year 2002
Program Actual Performance (for Continuous Improvement Model)	26%	34%	42%
State Average (for Relative Ranking Model)	34%	38%	38%

External Criteria Model

Although past performance is the basis for performance standard setting under the two models just discussed, policy considerations determine performance standards under the external criteria model. With this approach, policymakers typically set a long-term strategic goal as the standard that all programs must meet. Often the goal is set in response to a perceived crisis or political objective.

This model of setting standards is familiar to us in modern life. For example, automakers must meet fuel efficiency standards, measured by miles per gallon, and airports have standards for the percent of passengers and luggage that must be screened. In adult education, this approach often appears in the setting of GED pass rates or employment placements to improve performance. Several years ago, when the National Adult Literacy Survey (NALS) found that low levels of literacy were widespread, some states used this model for setting goals to raise literacy levels to a higher standard. For example, some states set standards to have all adults perform at least at Level 2 of NALS within 5 years.

As these examples show, the use of the external criteria model may be a good approach for mobilizing programs to adopt a long-term policy goal and achieve a uniform higher standard of performance. The approach can bring attention to the underlying problem, and mobilize resources to improve it, resulting in improvements. Fuel efficiency and baggage screening have certainly increased, for example. However, the chief drawback to the approach is that the external standard, by not considering past performance, may be set at an unrealistic or unachievable level. As a result, program staff may not take the approach seriously or may resort to counterproductive approaches to achieve the performance standard, resulting in unintended consequences (discussed later in this chapter).

Return on Investment Model

The purpose of this standard-setting model is to provide an indicator of whether the resources invested in a program are worth the outcomes achieved. Typically the monetary cost of the program is compared to the value of the outcomes. This model is common in business but is not often used in education or other social service programs.

Its most frequent use has been in employment training, where the cost of such programs is compared to the economic benefits that result from graduates' getting jobs. To determine the program's success, the increased taxes paid by graduates and reduced cost in welfare and unemployment benefits are weighed against the program cost.. The benefit of the GED credential also has been evaluated in this way.

At the program level, this model also can be used strategically to evaluate the value of classes or sites. For example, a program director can weigh the cost of running an instructional site by evaluating the outcomes it achieves compared to the resources it takes to operate. Under this standard-setting model, the director would calculate the cost of the program and set a value on the outcomes to determine the performance standard the sites need to meet to make it worth continuing.

Many people criticize this approach as cold and narrow and see this as its primary disadvantage. By using the cost-benefit approach inherent in this strategy, you may ignore other valuable outcomes the site or program produces. In the above example, the site may offer other services to the community or bring the community together in positive ways. Closing the site may leave the community with no services at all. These contributions are not reflected in the performance standards. However, the approach may be appropriate when cost considerations are important, particularly in an environment of cost cutting and performance pressures.

Adjusting Standards for Local Conditions

The performance standard-setting model you select represents a policy statement on the relationship between performance and quality that you want to instill in programs. Use of a continuous improvement model, for example, means you want the program quality to improve for every program, while the relative ranking model implies the policy of a more uniform level of quality across programs. External criteria can motivate programs toward a broad overarching goal, while the return on investment model embodies a cost-benefit approach.

Once you have set this general policy however, the next step is to consider whether the standard set this way will have the desired effect at the individual program level. Research on the effective use of performance standards demonstrates the importance of adjusting the standards for local conditions to make them work to improve program quality. Standards set at the wrong level will not work. They will either be too easy or too difficult for the program to meet. In either case, program improvement will not result.

It is not always necessary to adjust performance standards to local conditions. In fact, if you need to make too many local adjustments, it may mean you are using the wrong standard-setting model. Most of the time, only a few standards for a few programs need to be altered from the general model. There is no definitive way for deciding whether to adjust standards and how much to adjust. You must make these decisions by weighing policy considerations with local program conditions, along with performance

standard requirements placed on other programs. Three main factors affect program performance that may require you to adjust standards: (1) student characteristics, (2) local program elements, and (3) external conditions.

Student Characteristics

As in just about everything else in education, the most frequent refrain you will hear when setting performance standards is, "But my students are different!" The wide variation in student abilities, needs, and goals common to all educational settings, is especially challenging in adult education, where multilevel classes and very low-literate students are the norm. The usual problem is that the standard is too high for the program to meet because, unlike other programs, the program serves more lower-level students or students with unique problems. If this is true, the program will have difficulty achieving the standard and may try to "meet" it by such tactics as enrolling only higher-level students, placing students at a lower level so they can gain quickly, or manipulating the data.

The continuous improvement performance standard-setting model can mitigate this problem to some extent because, under that model, the standard is set based on the program's past performance. The range of educational functioning levels in the NRS also provides a way to adjust for student literacy levels, as you can set level advancement targets lower for students entering at the literacy or beginning levels. For example, you might set an advancement target at 10 percent for beginning literacy students and set 20 percent for the high–intermediate level.

The need to adjust for student characteristics may be greater with the other standard-setting models that do not inherently adjust for past performance. However, even with the continuous improvement model, three main conditions may make adjustments necessary.

- **An especially challenging group of students.** A program may serve students at an unusually low level of literacy, which may require a lowering of the overall standards. For example, English as a second language (ESL) students from cultures where there is no schooling or tradition of literacy, such as the Hmong or Somalis, may take a long time to achieve even the most basic literacy skills. Students with some types of disabilities also may progress much more slowly than other students. A program that serves a high proportion of very low-literate students also may have depressed rates of employment, GED passage, and entry into postsecondary education, requiring adjustments to the standards for these measures as well.

- **Students at the lower end of the level.** The educational functioning levels encompass a broad range of literacy skills. For example, beginning ESL encompasses two student performance levels (SPLs) on the Basic English Skills Test (BEST). When students enter class, their skills vary within the level, with some students at the bottom of the level, some near the top, and

others in between. The performance standard for a level is set as an overall average performance for the level. Adjustments may be needed, however, if programs serve a disproportionate number of students who fall in the lower end within the level.

- **An influx of different types of students.** A change in the program's student population is another reason why you may need to adjust standards. For ESL programs, this change may be a sudden increase in immigrants or refugee resettlements, for example. An employer layoff or closing of another educational program also may result in the enrollment of an historically different population into the program. This type of change in the student population may require an adjustment of performance standards from the number set through the general model you used.

While the above discussion has addressed the need to lower standards to accommodate serving hard-to-serve students, it is also possible to have students who make it *easier* for the program to meet standards. For example, the program may have students at higher levels of literacy who progress through educational levels faster and move on to community colleges, or the program have more job-ready or GED-ready students. In these situations, you may also want to adjust performance standards—but you will want to raise them.

Local Program Elements

Unique aspects of a particular program also may create a need to adjust performance standards from those set through the general model. Instructional arrangements, such as hours of instruction offered, instructional approach, and class schedules are the factors most likely to affect performance. You may need to use different standards to evaluate programs that differ significantly along these dimensions.

Due to resources, funding requirements, the types of student a program serves or its philosophy, programs differ in the amount of instruction they offer. For example, programs receiving welfare (TANF) funds may offer more intensive services, while programs serving a large proportion of working students may offer night classes that tend to be less intensive. Because the amount and type of instruction offered to students affects student learning and what they can achieve from class participation, you might expect better outcomes for longer, more intensive programs.

A program that uses different instructional approaches or programs that have more specific goals for instruction than other programs also may need special consideration. Examples of such programs include workplace literacy, where the focus is on work-related skills and Even Start or family literacy classes. If classes with these special goals create different student outcomes, it may be advantageous to adjust their performance levels. You might expect higher employment placements, but lower literacy gains for programs that emphasize workplace literacy, for example. The chief advantage

of adjusting performance levels is that you can create incentives for providers to offer these types of instructional services to implement policy priorities.

External Conditions

Every program operates in a community environment that reflects the types of students who attend and the services offered. These external conditions can also directly or indirectly affect program performance. The area's demographics, for example, influence who will attend, which in turn affects instructional approaches and student outcomes. As we have already discussed above, student literacy levels and other needs can make it necessary to adjust program performance levels. Because it is desirable that the program serve the community population, performance standard policies should be adjusted to promote and not inhibit serving diverse student groups and students with greatest need.

Another major factor external to the program that can affect outcomes is the community's unemployment level. If the program is located in an area of relatively high unemployment, for example, you may consider lowering the performance standards for entered and retained employment by some factor to reflect this problem. Conversely you might want to raise the standards for these measures in areas of low unemployment.

Transitioning students to further training, such as community college, is a core outcome in the NRS, as reflected in the NRS performance standards. In some areas, however, there may be few colleges or additional training opportunities to which students can advance. Setting the performance standards for this measure, then, requires a consideration of whether the programs in such communities can meet the standard. As with employment, if the achievement of the outcome is largely determined by this external condition, it is often advisable to adjust the standards for programs.

Other external conditions that can affect program performance result from unusual and often unforeseen circumstances that arise in an area. Circumstances, such as large layoffs by local employers, social service cuts, and natural disasters, might affect the availability of services, instruction, and student attendance. When these events occur, you should consider whether it is advisable or necessary to adjust standards for programs in the affected area to reflect the new realities.

To illustrate situations where adjustments to local performance standards are often necessary, the box below provides sample scenarios that may create the need for adjustments. Examples show the three most frequently used standard-setting models: continuous improvement, relative ranking, and external criteria.

Making Standards Work: Shared Accountability and Avoiding Unintended Effects

Selecting the standard-setting model appropriate to your state's policies and program realities and making adjustments as needed are central to an accountability system that promotes program improvement. However, just setting the standards is not enough.

A common mistake states make is to determine the performance measures and standards, inform programs of their accountability requirements, and then assume the system will result in good outcomes and program quality. No matter how well you set standards, you will not have effective accountability without a real system—a system that recognizes the need for shared accountability and technical assistance, includes appropriate use of rewards and sanctions, and carefully monitors performance.

Shared Accountability

Researchers and policy analysts who study educational reform and program performance agree that reciprocity between funder and grantee is fundamental to the success of accountability systems. The funder that awards resources can require accountability from the grantee, but it also has the responsibility to provide adequate resources for the grantee to understand and meet the requirements. The grantee that accepts the resources has the responsibility to meet the accountability requirements. With this sense of shared accountability, the state and its programs can enter into an agreement where the state provides the tools and environment for improved performance and the program agrees to work toward making improvements to meet performance requirements. If a sense of shared responsibility is lacking, programs may not be responsive because they will believe they are being held to unrealistic or unattainable standards.

Shared accountability begins through buy-in and collaboration with local programs in the standard-setting process. Staff in your local programs should have the following information:

- The purpose of the performance standards;
- The policy and programmatic goals the standards are meant to accomplish; and
- The standard-setting models you adopt.

Standards should be set in collaboration with programs, not simply imposed from above. In this way, you can gauge whether the standards are appropriate and realistic and whether adjustments or accommodations are necessary for local program conditions.

Adjusting Local Standards: Sample Scenarios

Continuous Improvement Model

Using a continuous improvement model, a state set performance standards for GED attainment at levels slightly higher than the previous year's for each local program. However in the previous year, several local programs received a grant to offer an extensive amount of "fast track GED" instruction prior to the release of GED 2002 and consequently, their secondary completion and GED rates soared. The "fast track" grant is now over and the program staff think the current levels set by the state are too high and should be lowered based on levels they attained before the grant.

Solution. Try to disaggregate the effect of fast track students from the program's past performance and review these programs' performance prior to the fast track grant. Adjust the standards appropriately based on these performance histories.

Relative Ranking Model

The state uses a relative ranking to set local performance standards. A local program that failed to meet its educational gain performance standards reviewed its student demographic data and finds that it serves a high proportion of older learners. Although the state average age of ABE learners is 33 years old, the local program's average student age is 49 years old. Program staff requests that the state adjust its standards lower, based on the common belief that older learners do not make gains as quickly.

Solution. Review learning gains by student ages across the state and in other programs located in the same area. If the analysis reveals that older students do in fact progress at a slower rate within the same educational levels as younger students, adjust performance standards based on expected older students' progress. For example, use the older students' average advancement when computing the relative ranking levels.

External Criteria Model

The legislature wants all employment and adult education programs to show at least a 20 percent increase in the percentage of participants who get jobs. In response, the state sets the standard for adult education programs for entered employment 25 percent higher. Several adult education programs claim that they cannot meet this standard because they serve a significant number of learners who are already working. Because these learners cannot set a goal of "obtain employment," the number of students with the goal is low, and the programs cannot substantially increase their entered employment rate.

Solution. Check these programs' data to verify the number and percent of learners who have an entered employment goal and who are employed, and the entered employment rate. Also determine the number of learners the program will need to place in jobs to increase their rate by the state-mandated 25 percent. Adjust the levels lower if you agree they cannot be met or provide technical assistance to assist the programs in developing strategies for placing the learners in jobs.

It is also important to collaborate on aspects of the program that are expected to change to improve program performance. For example, if the program is to improve its percentage of educational advancements, discuss with the program whether the expectation is that these increases might result from different instructional arrangements, better retention, or better assessment procedures. Mutual understanding about the reason for the standard and how the program can meet it creates a healthy environment that fosters accountability and program improvement. Technical assistance will help programs carry out needed reforms in their procedures to meet performance standards.

Accountability arrangements sometimes entail only the state creating standards and setting rewards and sanctions, expecting programs to comply or else. Such authoritarian relationships are prone to produce resistance and resentment, and when compounded by a lack of appropriate resources at the local level, are bound to fail. By providing resources and fostering communication, states can initiate a shared accountability approach that creates an atmosphere of trust where real change can occur.

Appropriate Rewards and Sanctions

Local performance standards must be set at the appropriate level. If they are too high, programs will not be able to meet them and quality will not improve. In fact, *unintended* effects are likely to result. If standards are set too low, program staff will not take them seriously because the standards can be met easily and again, there will be no program improvement. As discussed earlier, use of the appropriate standard-setting models with local adjustments as needed, will lead you to set standards at the optimal level. But while the level at which performance standards are set affects performance, their effectiveness also depends on the rewards and sanctions tied to them.

An accountability system for program improvement needs both a reward structure and sanctions to be effective. Programs reaching or exceeding their performance standards receive monetary or other rewards, while programs failing to reach their standards may be sanctioned by losing all or part of their funding. Research on accountability has shown that while both approaches are important, rewards are more effective in promoting program improvement than sanctions, all else being equal.

The use of sanctioning as a motivator for program improvement can be counterproductive, because the potential loss of funding puts tremendous pressure on a program to perform. By raising the stakes, programs see their survival at stake and will resort to any number of tactics—tactics that have nothing to do with program improvement—to meet their standards. Known in the performance management literature as "unintended consequences," these tactics range from enrolling only students at higher skill levels who are likely to improve and exit quickly, placing students in inappropriately lower levels so that they advance more quickly, playing statistical counting tricks where low performance is masked, or downright cheating. Indeed, in the public school system several notorious instances of dishonest reporting have recently come to light in response to intensive accountability pressures based on sanctions. Furthermore, overreliance on

sanctions gives an authoritarian aspect to accountability, which is counter to the need for shared accountability just discussed.

Reward structures, in which the program receives extra funding or similar bonuses for improved performance, tend to work better than sanctions for program improvement. Rewards do not create the pressures that sanctions do, because failure to achieve the reward does not typically affect the program's survival. Using rewards appeals to the program's sense of professionalism and desire to do well, which if well executed can create a healthy atmosphere of competition, as programs compete for recognition of a job well done. Setting rewards at the right levels, however, is important. When rewards are too high, they become the focus of performance, and unintended consequences may result, similar to the unintended effects of sanctions. If the rewards are too small, however, they will not motivate improved performance.

Although sanctions can produce unintended effects, some form of sanctioning is almost always needed in accountability systems, because rewards by themselves usually do not offer sufficient incentives, especially for low performing programs. There are two main ways to use sanctions that will minimize the likelihood of unintended effects. The method used most often is to set a probationary period for a low-performing program where it is given a reasonable time (a year or two) to improve. During that period, the state can help the program identify the problems and provide technical assistance to improve. Incremental goals can also be set to monitor improvement. Sanctioning and defunding of the program occurs only as a last resort.

A second way to minimize the unintended effects of sanctioning is to apply only a proportion of funding for program performance. For example, some percentage of a program's total funding is made contingent on good performance. With this approach, a program does not face total loss of funds if it performs poorly. As with reward structures, the state needs to set the level of sanctioning low enough to prevent unintended effects, but high enough for programs to take it seriously.

Data Monitoring and Local Contact

As with other complex processes, the artful balance of program performance and accountability requires vigilance to promote program improvement. Adult education programs are complex, and an accountability system has only an indirect effect on what programs do. Some of the effects of NRS requirements may be unforeseen or change in response to changing conditions in programs. The best way to know whether performance requirements are working as planned is to check data regularly and maintain good communication with your local programs.

Regular data monitoring, as we will discuss in chapter 4, will fulfill this purpose formally, and ongoing review of performance data will help identify problems, including unintended consequences, by alerting you to difficulties early on. Your data review should include examining, at least quarterly, performance standard measures and student demographics. By looking at these data over time, you can identify performance trends

that will help you understand what programs are accomplishing. Key measures to examine are as follows:

- The number and percentage of students who are pre- and posttested;
- Percentages and numbers of students who advance levels and attendance hours it took them to advance;
- Average attendance hours; and
- The number and percent of students who achieve NRS follow-up goals.

The data reports should be disaggregated by type of student and other pertinent variables; otherwise, the reason for the differences may be masked. The following patterns in these data will provide clues to possible unintended effects of performance standards or other data problems:

- A sudden change in performance over time, either higher or lower;
- Low variance in the data so that all students are scoring similarly;
- Large amount of missing data or *no* missing data; and
- Gaps in the data so that it clumps at low or high ends of the scale, rather than being more evenly distributed.

Such changes, if not explained by student or program changes, may signal data collection and reporting changes that have little to do with performance. It is also a good strategy to rely on your instinct, especially with programs you may know well. Sometimes the data just may not seem right. It is a good idea to continue exploring the data when you have this "gut" feeling.

Good, ongoing communication with your local program staff is also invaluable to making standards work. When you have a trusting relationship based on shared accountability, you can frankly discuss problems with meeting performance targets and work to resolve them. Regular contact with local programs, including frequent telephone contact, e-mail, a help line, or a Web site, and site visits will help foster better communication. State conferences provide an additional opportunity to discuss performance issues and their impact on local program performance.

Summary: Setting Local Performance Standards

Table 2–3 summarizes the steps for setting local performance standards presented in this chapter. The first decision is to select the standards-setting model: continuous improvement, relative ranking, external criteria, or return on investment. The model you choose should be the one that best reflects state policy goals for the program and that

Table 2–3
Summary of Local Performance Standard-Setting Process

Procedure	Process	Goal
1. Select standard-setting model.	Select one or combination of four models: • Continuous Improvement • Relative Ranking • External Criteria • Return on Investment.	Reflect state policies and promote program improvement.
2. Set rewards and sanctions policy.	Focus on rewards, set at moderate level; Set sanctions appropriately, allowing time for improvement.	Create incentives to meet standards while avoiding unintended consequences.
3. Review performance levels for local adjustment.	Assess local conditions, program practices, and student characteristics for each program.	Ensure standards are fair and realistic to help promote local improvement efforts.
4. Provide technical assistance.	Work collaboratively with program to identify needs to improve program performance and meet standards.	Create a collaborative atmosphere of shared accountability.
5. Monitor performance often.	Review data regularly and maintain ongoing, open communication with programs.	Identify and avoid potential problems that may hinder performance and program improvement and foster unintended consequences.

promotes local improvements in program quality. You may want to use different models for individual programs or some combination of models to meet these goals. For example, a continuous improvement model might be combined with a relative ranking model to give a program a short-term goal to improve performance over last year and a long-term goal to move closer to the state average (higher on relative ranking).

The state's next decision is to determine a system of rewards and sanctions tied to performance. The rewards and sanctions need to be high enough to motivate performance, but not so high as to create unintended consequences, where program performance becomes focused on meeting standards at all costs. In most cases, the optimal course is to set a moderate level of rewards, such as incentive funds for exceeding standards, and to set sanctions that are used as a last resort. It also is advisable to have a probationary period, in which low performing programs have an opportunity to receive technical assistance and make improvements, before you apply sanctions.

When setting standards, states often question how high to set them. It is essential to set standards at an appropriate level to motivate program performance and improvement without resulting in unintended consequences. When standards are too low, they fail to motivate programs, but when they are too high programs find them unrealistic and may ignore them—or programs may resort to desperate means to meet them. Unfortunately, there is no simple formula for determining the right performance level. The best approach is to consider the planned levels for individual programs in light of the local program variables and external conditions. Questions to ask when evaluating the levels you set for the standards include the following:

- Do you need to make accommodations for students or local conditions?
- Do the performance levels reflect state and local policies?
- Do the standards reflect the program's practices—do they make sense for what the program accomplishes and the types of students it serves?

After you consider these questions, you can then make adjustments as needed.

For accountability to promote program improvement, mutual responsibility for achieving performance outcomes is necessary. In what is often called shared or mutual accountability, local programs must meet performance standards, but states must provide resources to help programs make the changes needed to meet these requirements. States can provide resources through training and technical assistance. Good ongoing communication and ongoing data monitoring also are essential elements of successful accountability systems that foster program improvement.

This discussion of the performance standard process makes a critical assumption: that the data on which the system is based are of high quality—are valid and reliable and provide accurate measures of program performance. Without good data there can be no meaningful information on which to evaluate performance and no data-based planning for program improvement. If you have problems with particular data elements at some programs or for the state as a whole, you must address these problems. Another NRS resource, *Guide for Improving NRS Data Quality,* will help you identify and correct data-related problems.

Chapter 3. Under the Data: Performance Measures and Program Processes

Outcome-based measures such as those used in the NRS provide us with indicators of program performance. They can tell us what students are accomplishing as a result of their participation in the program. Such information is essential in these times of heightened program accountability, when demonstrated outcomes necessary for funding are part of the defining characteristics of quality programs. In this environment, administrators and educators dutifully report these measures, but may fail to reflect further on what the measures represent about how their program works.

Outcome measures are powerful and persuasive because they presumably reflect program quality. However, outcome-based accountability systems like the NRS include only measures of student performance and lack direct measures of program *process*—what programs actually do to produce these measures. Although it is assumed that good program processes produce good outcomes, this assumption needs to be verified through other means, including direct observation.

Often performance measures seem so removed from the program context that program staff does not explicitly recognize their connection to program quality. Data seems to be just numbers on a form, removed from reality. Yet, program improvement will not result from the accountability system unless the relationships between process and outcomes are understood. In this chapter we explore what is under the data—the complex set of programmatic processes and procedures that end up represented as numbers on forms.

Under the Data

The NRS outcome data you measure and report represent the end of a series of programmatic, instructional, and data collection decisions and procedures performed by many different people. The measure is the most directly observable product of this complex system under the data. To illustrate these underlying relationships, we present the NRS measures as "data pyramids." We put the measure at the top of the pyramid to illustrate that it is supported by a structure of processes and procedures upon which the measures are built. These processes consist of general administrative and leadership arrangements of the program, data collection procedures and program procedures.

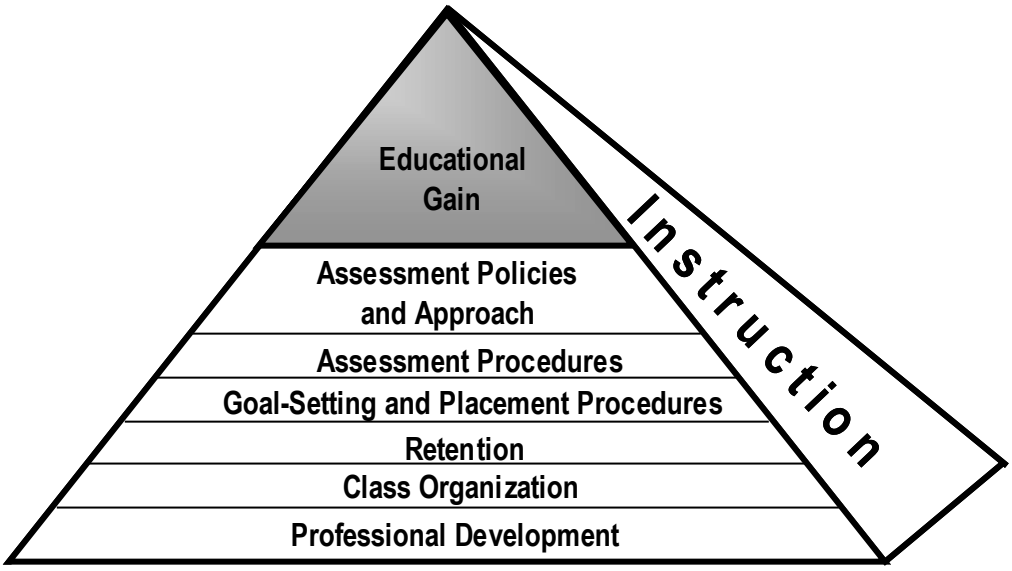
We present four sets of measure on the pyramids: educational gain; the NRS follow-up measures (entering and retaining employment, entering into postsecondary education , and obtaining a secondary credential); retention; and enrollment. Although the latter two measures are not NRS outcome measures, they have a direct relationship to outcomes, and many states set performance standards around them. In addition, a full understanding of program performance requires states to monitor these two measures and examine the program processes underneath them as well.

Educational Gain

The central measures of the NRS are educational gains. These measures represent program success in the most basic mission of adult education— to improve the literacy skills of participants. Education gain actually consists of 11 measures of the percent of advancement of students from one of the NRS levels to higher levels. Programs must make this assessment using a standardized assessment, with a pretest given just as instruction begins and a posttest given after specified hours of instruction. Students who are not pre- and posttested cannot be counted as having made gains.

Supporting the education gain measures are the program policies and procedures that are truly the heart and soul of an adult education program. All the major program policies and procedures come together to help determine this set of performance measures. Exhibit 3–1 illustrates the program processes that serve as the essential variables underlying the measures and affect what students learn, how they perform on the assessments, and how these measures are collected and reported.

Exhibit 3–1
NRS Data Pyramid: Educational Gain



Central to literacy development is the type of *instruction* offered by the program, which the exhibit shows as a supporting wall to the pyramid. Clearly, instruction has the greatest potential impact on educational gain and is the first place to look when examining program performance on these measures. Instruction intersects with all other aspects of the program operations. Effective instruction uses an organized curriculum based on research-based principles of adult learning and is responsive to student needs. At a minimum, good instruction needs to have the following characteristics:

- Incorporates content appropriate for the student skill levels, needs and interests;
- Is guided by a curriculum that identifies content standards, skills, or competencies;
- Is supported by related materials and activities; and
- Is of sufficient duration and intensity to affect learning.

Programs often have difficulty advancing students to higher educational functioning levels and often will be deficient in one or more of these areas. Likewise, high performing programs typically provide solid instruction that includes these features and serve as a model for improvement to low-performing programs.

Another variable affecting performance on educational gain measures is the program's *assessment policies and approach*. The NRS requires that every state have an assessment policy that specifies valid and reliable assessments, when and how programs are to pre- and posttest students, and how to match test scores to NRS levels. Local programs not only have to be aware of the policies but also need to implement them faithfully.

Valid and reliable measures of educational gain require that programs follow the proper *assessment procedures* for all tests. All standardized assessments have a formal protocol on how the test is to be administered. These procedures include the use of alternative forms, testing conditions, time allowed for the assessment, standardized scoring, and the amount of instructional time a student needs before posttesting. Violation of these procedures often invalidates the assessment.

Following state NRS assessment policy improves the accurate collection and reporting of educational gain, but assessment used only for reporting purposes is not sufficient to support the program's instructional approach and improve student learning gains. Effective programs also administer informal periodic assessments of student progress. These formative assessments may include alternative assessments and teacher- or program-developed tests. Assessment informs good instruction, serving as feedback to students and teachers that can be used to adjust instruction. Teachers can be learner-centered and can build on what students know and need only if they know what students have learned. A program with an effective approach toward assessment will have procedures to inform teachers of the results of both types of assessments and the implications of these results for instruction.

An effective program has an intake process that informs *student goal setting and placement*. This process allows the program to determine student needs, abilities, and interests and helps to identify the appropriate educational functioning level in which to place the student. Knowing student goals also guides teachers toward an instructional approach that is likely to be more interesting and effective for the student. Placement at

the right level gives the student the best opportunity for advancement. When students are placed at too high a level, they might become frustrated or bored, drop out, or not advance. Likewise, placing students at too low a level does them a disservice because instruction may be too easy and again, they may drop out.

Student *retention* is another critical variable in determining educational gain. Students cannot learn if they do not stay in the program long enough to improve their literacy skills. Furthermore, students who do not attend long enough cannot be posttested because a minimum amount of instruction is needed before tests can detect learning gains. Low rates of posttesting result in deflated performance on these measures. Consequently, an examination of program performance on educational gains must include a review of program efforts to enhance retention. We further review these factors below in our discussion on performance standards for retention.

The way a program organizes its classes can affect program performance on educational gains, as class arrangements influence both retention and instruction. *Class organization* includes how much instruction is provided and what educational functioning levels are offered. The number of hours of instruction offered by a program through its class schedule affects learning, as does the intensity of instruction. Programs may offer more intensive classes that meet for more hours for fewer weeks, while in other programs classes may be spread over a longer period but meet fewer hours at a time. The type of class that is most effective depends on several factors, including student needs, literacy levels and curriculum offered. The locations and times classes are offered also may affect student gains, as discussed later under retention.

Whether the program has multilevel classes where students are at similar literacy levels also affects the instructional environment and hence student learning gains. Although teaching multilevel classes is challenging, many programs must organize classes in this way due to limited resources or low enrollments at some levels or sites. Especially with low-level learners, however, multilevel classes may depress educational advancements.

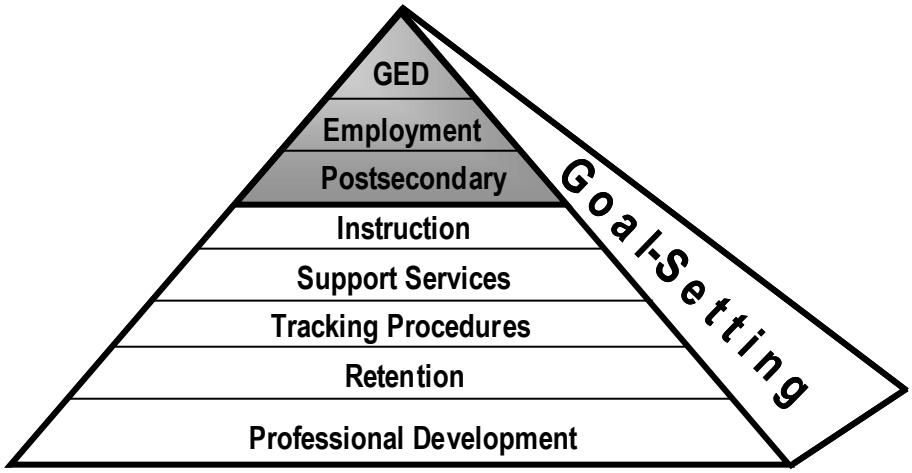
In the educational gains pyramid we have placed *professional development* as the foundation to emphasize its importance to all of the program factors we have discussed. A program may have exemplary policies, plans, and procedures for its operation, but they will be ineffective unless staff understands them and is trained on how to implement them. Teachers need ongoing professional development on instruction and assessment, and all staff should have training on NRS and state data collection requirements. In the previous chapter, we discussed the need for professional development for local staff to understand performance standards. Other training topics that will help program quality and performance include goal-setting and placement and methods for improving retention.

Follow-up Measures

The NRS includes four additional measures of student outcomes: entered employment, retained employment, entered into postsecondary education, and obtained a secondary credential. These measures are known as follow-up measures because programs do not report them until after the student exits and they require postenrollment data collection procedures. Although programs that serve adult secondary and other higher-level students have always had a GED component, reporting the employment and postsecondary entry measures were new to most states and programs within the NRS.

Unlike educational gains and GED attainment, where instruction, assessment, and other program activities have a direct effect on the outcome, some program components have a less direct influence on the employment and postsecondary outcome measures. Many factors besides the skills the student learns in the program affect whether a student gets and keeps a job, for example. Similarly, many nonprogram-related external factors affect student transition into community college. Despite these external factors, however, program practices can affect outcomes, and performance can be improved by directing attention to them. The data pyramid in exhibit 3–2 illustrates the main program variables affecting performance on follow-up measures.

Exhibit 3–2
NRS Data Pyramid: Follow-up Measures (GED, Employment, Postsecondary)



The supporting wall of this pyramid is the program's *goal-setting* procedures, emphasizing their key role in influencing this measure. When a student sets a follow-up goal, the program is held accountable for helping the student attain this goal during the program year for NRS purposes. For this reason, it is important to ensure that the goal for the NRS is realistic and that students are assigned to classes that will help them achieve their goals. Although setting a realistic goal is important for accountability purposes,

students' long-term goals should not be ignored simply because they are not attainable during the reporting period. Disregarding long-term goals not only does a disservice to students, but it also can harm your program by denying the opportunity to demonstrate that the program can help students achieve such goals.

An effective goal-setting process will help provide an accurate count of both the number of students who enter with follow-up goals and the number who achieve them. The process often includes a student orientation where students meet with counselors or instructors who help them set appropriate goals for their needs and skills. One effective strategy is for the student to set both immediate and long-term goals, with a realistic timeline for achieving them. For example, a student may have a goal to pass the GED tests, but function at such a low-literacy level that attaining this goal in the short term is unrealistic. By setting GED attainment as a long-term goal, the program can be responsive to learner needs without adversely affecting its accountability measures by setting a goal that cannot be met during the reporting period.

As with educational gains, the nature of *instruction* also can affect program performance on the follow-up measures. However, besides being of sufficient quality, duration, and intensity to affect student learning, the instruction should have some direct relationship to the skills students need to achieve the follow-up outcomes. GED preparation courses and adult secondary education (ASE) classes focusing on GED content areas are obvious examples. Instruction for students with employment goals might include teaching employment-related skills or literacy skills in an employment context, and classes with large number of students planning to enter community college could include some exposure to the skills needed to function in that setting, such as study skills and time management.

Because follow-up outcomes occur after the student leaves, the program loses some degree of control in helping the student achieve the goal. However, programs can increase the odds of student success by offering *supportive services* to students while they are still enrolled. Assistance with test-taking strategies for the GED is common in most programs. For employment, the program can offer employment counseling or job placement either directly or through another agency, such as One-Stop Centers. Such linkages are especially helpful for programs with a strong employment focus. Some programs have community college transition services where students who are ready to make this transition can receive information on enrollment procedures, tuition assistance, admissions tests requirements, and counseling related to college class selection.

Another program feature that affects performance on follow-up measures is the program's student *tracking procedures.* These procedures include a database with the ability to identify students by follow-up goal and an organized way to track students who are to be contacted later for surveys. For example, employment measures need quarterly tracking, requiring the database to report students with employment goals by quarter after they exit the program. If the state uses a data matching methodology, the program must collect Social Security numbers and if the local program conducts a survey, it must have a method for identifying students to survey and to record survey responses.

The data pyramid identifies *retention* as a program factor affecting performance on follow-up measures. As with educational gain, students need to remain in classes long enough to obtain the skills needed to achieve outcomes. Consequently, you should examine students' average hours of instruction and time in the program when you are evaluating performance on entered and retained employment, GED attainment, and transition to postsecondary education.

The pyramid also places programs' *professional development* as a foundation for performance on these measures. Staff needs sufficient ongoing training and support on goal-setting, instructional approaches, tracking methods, and the role of support services in promoting postprogram success.

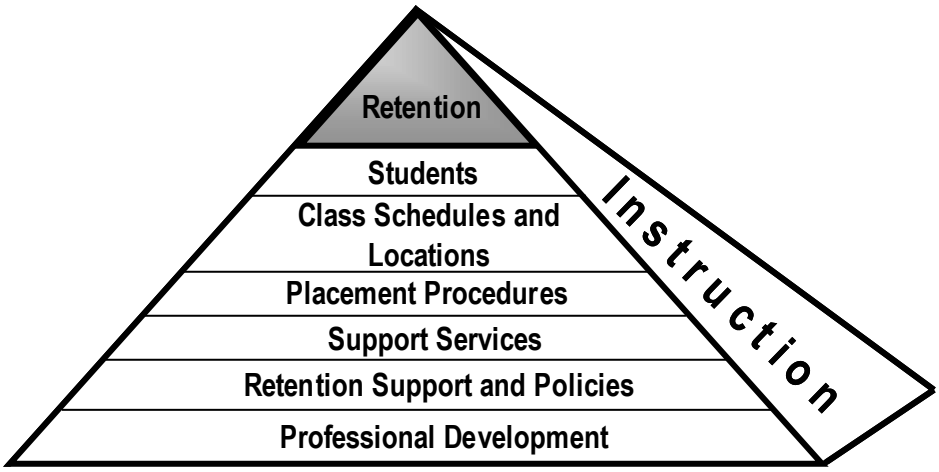
Retention

Student retention, measured by number of instructional hours or time in the program, is not an NRS outcome measure, and the U.S. Department of Education does not set state performance standards for it. Because of its importance in affecting outcomes, however, some states set retention performance standards for their local programs. These standards may be in the form of average hours of attendance or a minimum required percentage of students who persist for a set number of instructional hours. States set these standards on the assumption that a high student retention rate helps meet the Federal mandate to provide instruction of sufficient quality and duration for students to achieve literacy skills. As the data pyramids illustrate, we also have identified retention as one factor in program performance for educational gains and follow-up measures.

Exhibit 3–3 shows the data pyramid for retention, identifying seven program processes that can affect performance on this measure. We have placed *instruction* as a supporting wall for retention, on the assumption that it is the main factor that helps retain students in the program. Quality instruction that meets student needs will maintain students' interest and increase the chances that they will remain in class. In addition to the other characteristics of quality instruction already discussed, the relevance of the instruction to students, the delivery of instruction in an engaging way and with content at an appropriate level of difficulty, will promote retention. When students do not think they are learning or do not need to know what is being taught, the instructional approach will not hold their interest and they are unlikely to persist.

Adult education is unique among education and training programs in the wide variety of students who attend. Students vary by literacy level, native language, reasons for attending, age, cultural background, and a host of other variables. Not surprising, these *student* differences translate into different retention and attendance patterns. For example, employed students who attend at night typically attend for a much shorter time than other students. National data show that ESL students attend for more hours on average than ABE students, and recent research studies have identified retention patterns related to several types of student groups. To understand program performance on retention, attendance must be disaggregated by type of student group and types of classes.

Exhibit 3–3
NRS Data Pyramid: Retention



Because student needs vary, programs' *class schedules and locations* affect retention. Because adult education programs are voluntary, classes that are convenient to students will reduce barriers to attending and may positively affect retention. Location is critical, especially if students must rely on public transportation. The needs and situations of students in the program should inform scheduling of class times as well, to the extent resources allow. Offering classes in students' neighborhoods at class times that match their needs will increase retention and improve overall program performance on outcome measures.

Student engagement in instruction is difficult when the instructional content is either too easy or too difficult for students. If students become bored or discouraged, they may leave. Thus, a program's *placement procedures* give an indication of the program's ability to retain students. Indicators of good placement include the assessment methods used to identify student educational functioning level, effective student orientation, and limited use of multilevel classes.

Assessments used for placement not only should give a valid indication of student literacy level, but also should address the skill areas that match the planned instruction. For example, using an English grammar test to place ESL students in classes that focus on oral language may not produce the right information for placement. Some sort of orientation or trial period before placement is finalized will also promote accurate student placement at the right educational functioning level. Although multilevel classes are the norm in many adult education programs due to limited resources, instruction is difficult when students' literacy levels are too diverse within the class or when a program has too many such classes. Maintaining student interest and motivation is challenging when student skill levels vary widely and low retention may result.

Programs can increase their performance on retention by helping students overcome some of the external barriers to attendance. Adults' conflicting schedules and job demands, along with home, transportation, and childcare problems are the most common barriers to participation. Programs can help students improve attendance by offering them *support services*. Unfortunately, many programs lack resources to provide support services directly. However, if your program has a systematic way of identifying student needs, you may be able to help students overcome barriers to attendance through referral or problem solving and thereby increase students' resolve to attend.

One approach many programs use to improve retention is *retention support*. These support procedures improve retention by identifying students who attend sporadically or who are absent for extended periods. Teachers or other staff contact such students to identify reasons for poor attendance and help students resolve issues. Some programs also have *retention policies* or attendance requirements for students to remain in class. For example, after several consecutive unexcused absences, students may be dropped from class. Similarly, programs have adopted managed enrollment policies to promote attendance. In contrast to open entry/open exit policies, students may enroll in classes only at certain times, such as within 3 weeks of the class start date. Such policies establish an expectation among students that attendance is required and important, which often improves retention.

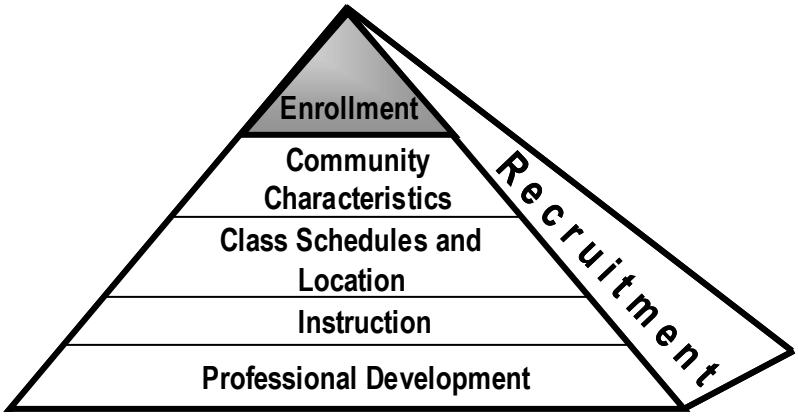
As in the other pyramids, *professional development* is a foundation of good retention. Teachers need meaningful training on all aspects of the program's retention policies and procedures, such as assessment and student placement procedures, the role of learner-centered instruction, attendance policies, and support services.

Enrollment

Enrollment—as measured by the number and characteristics of students who enroll in the program—is not an outcome measure, and the NRS has no enrollment standards. However, many states set enrollment targets or performance standards, which can include total number of students and proportions of students with particular characteristics. For example, the state may require a program to have a percentage of students at low-literacy levels or diverse ethnic populations. Other common enrollment requirements are for programs to have enrollment that matches community demographics and a balanced proportion of ABE, ESL, and ASE students.

The types of students a program serves greatly affect all aspects of its policies and procedures. Consequently, an examination of *recruitment* practices is informative of overall program quality and performance on NRS outcome measures. Exhibit 3–4 shows recruitment as the supporting wall or key variable in program enrollment. Community characteristics, class schedules and locations also affect recruitment and enrollment.

Exhibit 3–4
NRS Data Pyramid: Enrollment



Because the demand for instruction often exceeds availability, many programs do not need to recruit students, and the students who do show up are sufficient to meet program enrollment requirements. Even programs with large enrollments, however, sometimes may need to reach out to special populations, such as students with the greatest needs or a more diverse group of students. The program may have too few low-literate students to meet its grant requirements, for example, or may be failing to enroll different immigrant groups. In these cases, programs must engage in more active recruitment.

In addition to the usual flyers and advertisements, programs can rely on student networks and community organizations to recruit potential students. Often students enroll in the same program that their friends, family members, and spouses attend, so making your program well known in the community will help recruitment efforts. Targeting the message to specific student populations, as well as appealing to student needs that can be met by the program, are often effective recruitment methods.

Effective recruitment also requires knowledge of *community characteristics*, including the demographics of potential students. For example, recruitment strategies for ESL classes that target Latinos cannot be expected to work in a community of Asian immigrants. Many programs' enrollment goals include serving a representative proportion of adults in the community. Therefore an understanding of the community through Census data and needs assessment, for example, will guide recruitment and enrollment efforts.

Program enrollment also is dependent on the program's appeal to potential students through the *class schedules and locations* it offers. As with retention, convenient locations and times that meet the needs of busy adults will lower external barriers to participation. Schedules and locations that meet community needs will also positively affect enrollment, again underscoring the need to understand community characteristics to manage and improve enrollment. A community with young, employed adults who are

largely immigrants, for example, will want different class schedules than will older retirees who have lived in the United States for many years.

Similarly, the type of *instruction* the program offers also should match needs of students in the community to affect enrollment. Instructional factors include whether the program has the right balance of ABE and ESL offerings at the appropriate literacy levels and instructional content that meets student needs.

Finally, *professional development* is again a foundation to enrollment efforts. Staff and teachers who understand program enrollment needs and community characteristics, as well as the relationship of enrollment to instruction, will help improve program performance.

Quality Indicators as Program Standards

Standards for effective performance provide a systematic way to determine whether programs are performing in a way that promotes program quality. Over the last several years, there have been many state and Federal efforts to define effective practice and develop program standards tied to outcomes. One set of program standards that is widely used, known as quality indicators, provides an organized way to evaluate program performance that many states find helpful.

In the early 1990s the National Literacy Act (NLA) required states to develop model indicators of program quality for adult education. Using the model standards developed by the OVAE, most states developed these quality indicators to define good program policy and practice. Many states continue to use these indicators to guide their program monitoring and program improvement efforts. In the context of NRS outcomes and their relationship to program processes, states may consider using something like the quality indicators to set criteria on what constitutes a quality program and tie these standards into the accountability systems. Indeed, several states already do this as part of their approach to program monitoring.

As defined by the NLA, quality indicators measure "efficient and effective performance" of adult education programs. The model indicators cover educational gains, recruitment and retention, and include the program process areas of support services, staff development, curriculum and instruction, and program planning. The model indicators include examples of ways to measure them. For example, the indicators and measures for curriculum and instruction and professional development are as follows:

- **Curriculum and Instruction Indicator:** The program has curriculum and instruction geared to individual student learning styles and levels of student needs.

  – *Measures:* Instructional content addresses individual student needs (measured by classroom observation or self-report).

  - Student goal-setting processes are linked to instructional decisions.

- **Professional Development Indicator:** The program has an ongoing staff development process that considers the specific needs of its staff, offers training in the skills necessary to provide quality instruction, and includes opportunities for practice and systematic follow-up.

  - *Measures:* Staff development based on research-based practices;

  - Formal process to identify staff development needs.

Defining the measures and standards along with the indicators provides states with a uniform set of standards that define a quality adult education program. Quality indicators provide a systematic way to relate program process to outcomes. We return to a discussion of the use of quality indicators in the next chapter.

In this chapter we have taken a brief look under performance data to identify the program variables that affect the data. Performance standards and other NRS data provide a window on program performance and offer clues as to which program variables are working and which need improvement. An understanding of these relationships can guide program monitoring for program improvement. We now turn to a discussion of monitoring itself. Program monitoring allows a formal way to assess this connection between outcomes and process, as we will more fully explore in chapter 4.

Chapter 4. Planning and Implementing Program Monitoring

How often have you heard the terms *monitoring* and *evaluation* used in the same sentence? What was your immediate reaction when you heard these terms? For some people and programs, words that come to mind might include *stress, ranking, grading, sanctions*. Yet, most of us probably have been part of some positive evaluation processes. For example, performance appraisals or annual reviews can be helpful when we assess our own performance, discuss our accomplishments, receive feedback, and then set goals for the upcoming year. Although stressful, these reviews offer an opportunity to "take stock" and to think about where we would like to be in the future and how we might get there. Similarly, program monitoring and evaluation are opportunities for local programs to take stock and think about changes in their programs.

In chapter 2 we explored the use of the performance standard-setting process as a tool for program improvement. In chapter 3 we looked under the data to identify the program factors that affect performance. In this chapter on local program monitoring, we explore the methods for effective integration of quantitative and qualitative data while monitoring. We will expand on the use of NRS and local data for program improvement and return to indicators of program quality as tools for local monitoring and program strengthening. We will consider local program monitoring as an ongoing process of reviewing and documenting program outcomes and processes to identify and plan program improvements and enhance the program's capacity to meet the performance standards set with the state.

First, we will define the value of local program monitoring, and then we will look at approaches and steps for setting up an effective monitoring system. We will share sample tools and strategies that can be reviewed and adapted to meet state needs. States differ a great deal in terms of available resources for local program monitoring, including staff size, the number of programs, budgets, priorities, and regulations. We hope the ideas, tools, and strategies will stimulate adult educators to rethink the role of local program monitoring as a process leading to continuous program improvement.

The Value of Monitoring

Whenever we talk about accountability systems, someone will eventually mention one of the following refrains:

"What is counted becomes what counts."
"We treasure what we measure."

In designing local program monitoring, we want to explore and clarify what we treasure and find ways to measure it effectively so that we are indeed counting what counts. We look at performance results over time, we consider both process and progress indicators, we identify areas of strength and weaknesses, and we help programs make

informed decisions for program improvement. Therefore, local program monitoring should be considered an interactive, ongoing process, rather than a single event.

To monitor effectively, we need tools and strategies that are clear, and we need to include multiple stakeholders throughout the process. Regular monitoring reviews benefit local programs when we do the following:

- Keep local program staff focused on outcomes and processes;
- Identify issues of importance to staff and program participants;
- Increase staff involvement in the process;
- Provide data on accomplishments (positive outcomes and strong processes);
- Help refine data collection processes and products;
- Identify areas for program improvement;
- Provide information for decision-making; and
- Allow for enhanced program accountability.

We must simultaneously recognize, however, that monitoring is not an easy task for the following reasons:

- It can be difficult for programs to recognize the link between outcomes and processes. (How can we demonstrate that program outcomes are really tied to what happened in the educational program?)

- It is difficult to develop criteria for assessing program effectiveness when some learner outcomes are not easily quantifiable. (How can we document the qualitative changes in learners' lives?)

- Programs and staff fear judgments in high-stakes environments. (How can performance-based funding, incentives, and sanctions be fully understood by stakeholders?)

Approaches to Monitoring

Accountability requires monitoring, documenting, and certifying the integrity of data that are gathered from local programs and then aggregated at the state level. For some states, these accountability measures have increased the monitoring burden in the midst of diminished state funding. However, states now have clearly defined accountability measures, concrete baseline performance data (2001–03), and the capacity to gather data electronically. In essence, states now have the opportunity to rethink their monitoring plans and strategies to take full advantage of the accountability systems currently in place.

One solution that states are adopting is to include both desk reviews and onsite reviews. An onsite review is usually a single event lasting 1–3 days, while a desk review is an ongoing process. NRS data allows monitoring of programs throughout the year via

desk reviews instead of limiting local program monitoring to onsite reviews for a certain number or a certain percent of programs each year.

Desk Reviews

Desk reviews are especially useful for monitoring quantitative data. They provide a structured way to look at information that local programs already submit to the state office such as proposals, performance measures, program improvement plans, schedules, staffing patterns, and budgets. State staff can design monitoring tools for reviewing and providing feedback on reports that programs submit regularly (e.g., monthly, quarterly, midterm, or annual reports). Without significantly increasing the workload for state staff, desk reviews can guide staff in monitoring program activities and progress. With desk reviews, staff communicate regularly with local programs regarding anticipated changes or needs for technical assistance.

The findings of desk reviews can be used to focus the more intensive onsite reviews. Table 4–1 offers other advantages and disadvantages of desk reviews for monitoring.

Table 4–1
Advantages and Disadvantages of Desk Reviews

Advantages	Disadvantages
Data, reports, proposals, and plans are usually available in the state office.	Assumes that data are accurate and that reports and proposals reflect what is actually happening at the program level.
Reviews can be built into regular workloads over time.	Might not include voice/perspectives of multiple local staff members.
Data are often quantitative – can be compared to previous years' data or to state standards.	Gives static view of data and program process instead of dynamic interaction in context.
No travel required: no time out of the office and no travel expenses.	Lacks "team perspective" and opportunity for capacity building if not followed by an onsite review.

Onsite Reviews

Onsite reviews allow states to verify data and to look at the processes and qualitative information about a local program. Onsite reviews provide an opportunity for gathering more data and for elaborating on the initial findings of desk reviews. They allow for in-depth discussions with staff related to the program quality indicators and for meaningful interviews and observations with a range of staff members, learners, and program partners. Equally critical in these times of accountability, onsite reviews allow monitoring teams to access files and documents to verify their clarity, integrity, and appropriateness. While on site, team members look for recurring themes and patterns as they emerge from the document reviews, interviews, observations, and the discussions related to the indicators of program quality and to program data. Table 4–2 shows some of the advantages and disadvantages of onsite reviews for monitoring local programs.

Table 4–2
Advantages and Disadvantages of Onsite Reviews

Advantages	Disadvantages
Opportunity to look at process and program quality factors and to gather qualitative data. Teams can verify proposed plans and can check data quality.	Stress for local program and team to "look good."
Input from different perspectives (e.g., interviews and observations). Diverse team members can offer input.	Limited time frame for gathering information, interviewing, and observing. Setting up and training teams takes time.
Visibility in the field. Staff can explore options for improvement and provide technical assistance.	Requires time away from office.
Capacity-building. Other team members learn new strategies and share their expertise.	Travel time and expenses.
Opportunity to acknowledge strengths, offer praise, and identify best practices in the field for program-to-program mentoring.	

Before starting onsite reviews, team members need to fully understand their roles and how to use the monitoring tools with the criteria specified by the program quality indicators. For example, when observing a class or orientation session, monitors need to note what happens as objectively as possible based on the indicators. Their role is not to evaluate the teacher or presenter during this observation. Interviewers should use established protocols and probe gently to clarify, but not judge, the value of what is being said. Similarly, documents should be reviewed based on criteria stated in the quality indicators. For example, when reviewing individual learning plans or professional development plans, reviewers look for evidence specified in the indicators.

Ways To Collect Data for Monitoring

Standard yet flexible monitoring tools can effectively guide staff and monitoring teams to look at both quantitative and qualitative data from a distance with desk reviews, and up close in person with onsite reviews. Table 4–3 provides an outline of the main strategies for collecting data for program monitoring: program self-reviews, document reviews, observations, and interviews.

Program Self-Review

Prior to a monitoring visit, local program staff can complete a program self-review (PSR) based on program standards or other indicators of program quality. Later, the results of the PSR can help guide the program improvement process or help focus an onsite review. The PSR might reveal specific strengths of a local program that can be models for other programs. Or, the PSR might reveal indicators that require more discussion and observation during the onsite review. PSR findings are useful for helping to frame the written report. Depending on size, dispersion, and diversity of the program, the PSR might be completed by a sampling or the entire local program staff.

Table 4–3
Data Collection Strategies for Monitoring

Methods	Who or What	How
Program Self-Review	Staff	Review indicators of program quality .
Document Reviews	• Learner records (forms, test scores, attendance, etc.) • Learning plans (goals and work plans) • Portfolios • Curricula/materials • Recruitment materials • Promotional information • Referral information • Budget and financial reports	Conduct environmental scan— • Resource room, • Postings, bulletin boards, and • Flyers/Brochures for pick up. Use quality indicators to review for completeness— • Learners' files, plans, etc.; • Staff records, lesson plans; • Meeting notes; and • Curricula and materials.
Observations	• Classes • Intake and orientation • Learning lab • Meetings	• Take notes during observation. • Complete a checklist. • Tape or video record.
Interviews	• Learners • Staff • Partners and community leaders • Nonparticipants	Conduct open-ended or structured interviews: • One-on-one, and • Focus groups or classes.

Document Reviews and Environmental Scans

Many program documents can be reviewed from a distance, such as proposals, qualitative and quantitative data reports, and improvement plans. It is during onsite reviews, however, that the monitoring team can see if the documents are available and are being used effectively. For example, if teachers are not using a curriculum, program staff needs to explore why the curriculum is not being used, and then create a plan to encourage curriculum use, offer professional development opportunities, or, if necessary, begin revising or developing new curricula.

One of the main goals of the onsite review is to verify the accuracy of program data. The review team should audit a random sample of student files to verify that the print information matches the database. Local programs should include the following in student files to allow a review of data quality:

- Student identification and demographic information;
- Attendance rates;
- Years of schooling and placement level at program entry;
- Initial learning goals;
- Specified pre- and posttesting student information; and
- Entry and update records.

The review should compare this data with a print out of the same data from the program's database.

Observations

In education, observation is often the primary tool for teacher evaluation and, as such, may be overlooked when monitors think of data collection. For accountability purposes, states might consider observation in a way that is more akin to bird-watching—looking and listening in specific environments to identify certain features and behaviors. During an onsite review, monitors can observe interactions during meetings, at intake or orientation, in the hallways, in the computer lab, and in the classroom.

Skilled observers usually identify the activities or behavior they expect to see. For example, the program self-review might have cited strengths in using curriculum to guide instruction for life skills. Observers will look and listen specifically for instruction based on learners' lives and for the integration of authentic materials. They listen and observe while suspending judgment. Some take notes while others record what happens. The box below provides a situation and a set of questions to guide an observer as a program monitor.

Observing With the Monitor's Eye

Imagine you are sitting in a room where 15 students are studying fractions and taking turns going to the board to solve problems from their textbooks. The instructor comments on each problem and asks students to identify where they might use fractions in daily life.

If you are looking for indictors of instruction based on learner goals and needs, you might want to note:

- ✓ How many students are actively participating? Are they attentive when others solve problems or speak?
- ✓ How are students reacting to the lesson? Do they show interest/boredom?
- ✓ How do students demonstrate understanding or confusion?
- ✓ Are students able to relate fractions to their lives?
- ✓ Is there student-to-student interaction (peers/small groups)?
- ✓ Do students initiate questions beyond the textbook and lesson?
- ✓ Is there a variety of teaching strategies and activities?

Immediately following observations, monitors might review their notes to identify the links with the (a) indicators of quality, (b) activities in the program plan, or (c) recent professional development workshops. Later, monitors might prepare bulleted notes for the onsite debriefing, which can be included in the written report to qualify or support the recommendations and commendations. Using notes from observations can help illustrate the data cited in the report.

Interviews

Interviews provide rich data about people's opinions, knowledge, and needs. Although interviews are time-consuming, they can provide additional, first-hand information about what has been learned through the program self-review, document reviews, and observations. Interviews can help clarify information, explore ambiguous findings, or even provide contradictory information.

Who are the interviewers and the interviewees? Because onsite reviews strive to capture a great deal of information in a short period, it is essential to carefully select the interview questions for students, staff, and partners. Consider interviewing staff at various levels since administrative, instructional, and support staff might provide different perspectives on the various program components. Community partners or collaborators might know something about learner needs that are not known to the adult education program staff (e.g., higher education partner might be able to provide information related to enrollment opportunities for advanced ASE students). A social service provider might know about childcare services for learners participating in adult education classes. Interviewing employment service providers can open doors for opportunities such as job training, additional classroom space, job referrals, or guest speakers on workplace readiness and interviewing.

Learners at different levels or in different program components (e.g., ABE, ASE, ESL, Family Literacy, Corrections) might have very different experiences within the program and might be more or less vocal. Native-born and foreign-born students reflect different cultural perspectives and experiences. As a result, their willingness to be candid might require different interviewing skills. When monitoring a program that serves a significant ESL population, the team might consider using bilingual interviewers, if possible.

How are interviews conducted? Depending on the available time and human resources for interviews, teams might conduct some interviews with individuals and others with small focus groups. Before beginning, consider how the interview information will be used and shared. Open-ended interviews may seem ideal for gathering a full range of perspectives, but they are more difficult to synthesize for debriefing and for preparing a final written report. Similarly, some monitoring team members may not be as comfortable conducting interviews as others. Many states have found it helpful to have a structured interview protocol to ensure is consistency across team members and across programs. Interviewers need to listen carefully and note what is said, recording direct quotes whenever possible. Quotes from students, teachers, or partnering agencies can help enliven the written report and illustrate the quantitative data findings.

Interview questions should reflect the aspects of program quality that affect the NRS measures. The data pyramids in chapter 3, as well as the state's program quality indicators, suggest the topics you might address. Below we list sample questions for learner, staff and partner interviews.

Learner Interviews

1. How did you hear about the program?
2. How has the program helped you reach your goals?
3. Have you been able to attend on a regular basis? Why or Why not?
4. Were there particular ways of providing instruction that helped you? Have you learned how to learn?
5. What do you tell other people about the program?

Staff Interviews

1. Is there a staff handbook? Are you aware of the program's mission, goals, policies, and procedures?
2. How do you address the different learning styles of adult learners? How do you work with multilevel groups?
3. Do you have a role in developing curriculum, testing, intake and goal-setting?
4. Do you receive paid time for preparation and professional development?
5. Have you engaged in peer-coaching activities or worked on projects in collaboration with other staff?

Partner Interviews

1. How would you describe your relationship with the adult education program?
2. What do you think are the most critical needs for adult learners in this community?
3. Do you have any ideas or suggestions for improving the exchange of information and referrals?
4. Do your clients participate in adult education classes? If not, what do you think are the barriers to participation?

States can maximize monitoring opportunities by developing a monitoring plan supported with standardized tools and strategies that can be used consistently. The strength of a state's monitoring plan will be evaluated by its ability to use qualitative and quantitative data to document a program's strengths and weaknesses and to prepare recommendations for program improvement. We have included sample monitoring strategies from two states at the end of this chapter.

Monitoring With Indicators of Program Quality and Data

We presented NRS data pyramids in chapter 3 to illustrate the program variables that influence performance, as measured through the NRS. We also discussed how these program measures are embodied in the indicators of program quality or other program standards that states have developed to evaluate local programs. Table 4–4 provides examples of how quality indicators can guide the monitoring process. Column 1 identifies program areas addressed through the indicators of program quality. Column 2 suggests NRS data sources for each indicator, and column 3 outlines questions that can stimulate exploration during program monitoring. Column 4 provides examples of review activities that can be performed during a desk review, as well as those that can be conducted during an onsite review. Column 5 provides some sample findings related to the indicators and the reviews.

For example, a program quality indicator for management usually includes data reporting. Our primary source when monitoring a program's implementation of reporting is the database for collecting NRS data. A monitor can verify if the local program is using the data to enter data accurately and efficiently. During a desk review, staff can scan the database to see if learner information is being entered and updated on a regular basis. A data scan might also reveal if the program is serving the number of learners at the levels projected in their proposal. Later, during an onsite review, a monitoring team might conduct a random file review to verify that the information in the database (a) reflects the information provided by the learners at intake, and (b) matches placement and testing information.

In this example, the initial desk review indicated that few learners had been entered into the system. During the onsite review, the local administrator explained that teachers were holding the learner intake and update forms throughout the instructional term. During interviews, the data entry staff voiced frustration about hundreds of tattered forms arriving in folders at the end of cycle and stated that it was often impossible to gather missing information from teachers and learners after the course had ended.

As you scan the chart, remember that these are samples taken from different programs at different times. Monitors need to be flexible in discovering patterns or trends, which may not be captured by a single indicator or in the data. Onsite reviews allow for further exploration, and they can reveal strengths or weaknesses that might not be noted in a desk review.

Table 4–4
Monitoring With Indicators of Program Quality and Data

Program Areas in Indicators of Program Quality	NRS Data Source	To Explore/Questions	How		Sample Findings
			Desk Audit	Onsite	
Program Management (Data Reporting)	Data collected, reported, and entered into NRS database	**Are data valid and reliable? Is information being collected and entered accurately and in a timely fashion?**	***Scan*** Database for updates	***Review*** Learner files	Teachers holding learner forms throughout the term. Data-entry problems exist because learner information is incomplete.
Program Management (Collaborations)		Are there working linkages with employment and postsecondary providers?	***Review*** MOUs or Agreements	***Interview*** Partners	Active referrals between adult education, Workforce Investment Board, and vocational training.
Recruitment	The number and percent of learners enrolled in different programs	Does enrollment match demographic needs? Is the program serving the number of learners in proportion to need documented by Census? • Non-high school completers • Influx of foreign-born	***Review*** NRS Table 1, Census data, Increased K–12 ESL enrollment	***Interview*** • Learners • Community leaders • Teachers • Intake staff	Local program with high rate of non-high school completers serving only 13% in ABE/ASE. Immigrants represent 12% of population but comprise 55% of enrollments.
Goal Setting (Intake and Orientation)	Student data and goals collected; information entered in database	Is the intake and orientation learner friendly? Are student goals accurately identified and tracked?	***Review*** NRS Tables 5, 8, 9, and 11 for number of learners by goal and goal attainment, relative to overall enrollment	***Interview*** Learners and staff ***Observe*** • Intake • Orientation	Daytime classes: • Special orientation offered every week by coordinator. • Higher persistence rate. • Instructional quality varied in classes observed. Evening classes: • Orientation given in class by teacher. • Lower persistence rate. • High-quality instruction observed in all classes.

Table 4–4
Monitoring With Indicators of Program Quality and Data (Continued)

Program Areas in Indicators of Program Quality	NRS Data Source	To Explore/Questions	How		Sample Findings
			Desk Audit	Onsite	
Educational Gains	Number and percent of learners advancing by type and level	How do gains differ by— • Levels within ABE, ASE, and ESL? • Location? • Schedule- intensive vs. "non"? • Teacher or class? • Demographics (ethnicity, age, education background)?	***Review*** NRS Tables 4 and 1 Proposal (schedules, locations, staff)	***Observe*** At different sites • Trends? • Patterns? • Notables? ***Interviews*** Teachers Students Site coordinators	Some ASE teachers rely solely on textbooks and state that gains are low due to student retention. One ESL teacher corrects every mistake, provides no "wait time," and talks 75% of class time. In interviews, students talk about how good the "other" teacher is.
Assessment	Placement levels, and pre- and posttest dates with test scores; and pre- and post-matches	How can program increase the percentage of pre- and post-matches? Does assessment inform instruction and monitor individual progress? Disaggregate data to determine possible reasons for differences.	***Review*** Local data, class schedules, and other data (employment, children K–12)	***Discuss*** With staff	ABE/ASE classes are mostly during the day (4 times per week) ESL is offered 2 evenings per week. More ESL learners are employed and have school-aged children.
Curriculum and Instruction	Number of hours of instruction and number of learners advancing a level	Do all teachers have a curriculum to guide instruction? Does the curriculum and instruction address learner needs and styles, and offer variety?	***Review*** Proposal, reports, improvement plan, and curriculum	***Interview*** All staff and learners ***Observe*** Classes	Program is waiting for state content standards before developing curriculum. Students say that teachers use "real-life" materials sometimes.

Table 4–4
Monitoring With Indicators of Program Quality and Data (Continued)

Program Areas in Indicators of Program Quality	NRS Data Source	To Explore/Questions	How		Sample Findings
			Desk Audit	Onsite	
Persistence	The number and percent of learners completing level and leaving program before completing level	Does the program offer effective schedules, instruction, and locations? Does it have a plan to promote persistence? Review disaggregated data by– • Location or teacher; • Schedule (A.M. vs. P.M.); • Demographics (ethnicity, age, education background).	***Review*** NRS Table 4 crossed with Table 1 data	***Observe*** Different sites and times ***Interview*** Teachers and students in class with high drop out rate	Morning classes have more females and onsite childcare. Evening class students cite lack of transportation as barrier to attendance. Absenteeism leads to "falling too far behind."
Support Services	The number and percent of learners leaving program before post-testing	Does the program offer support systems to promote engagement and learning?	***Review*** NRS Tables 6 and 11	***Interview*** students, intake counselor and social service partners	Referrals consist of handout given to learners. No direct counseling. No systematic follow-up on "making connections" or "navigating systems."
Professional Development		Does the program offer professional development opportunities? What are the outcomes of professional development? Is the program getting a good return on investment?	***Read*** Proposal/Reports: • Professional Development (PD) Plan • Qualifications ***Request*** Information on most recent workshop	***Review*** Agenda and session designs for recent workshops ***Observe*** Classes and ***Interview*** Teachers who had participated in PD	No project-based learning was observed. Teachers say class schedules of 2 nights for 2 hours are insufficient for project-based learning.

Seven Steps for Developing a Monitoring Plan

Good monitoring requires careful planning if it is to be effective in identifying potential problems and promoting program improvement. The planning begins with getting support and buy-in from all stakeholders in the process, including local programs. The state must also decide on the scope of monitoring activities, given state policies and the resources it can devote to this important activity. At this point, you will want to consider the advantages and disadvantages of desk reviews and onsite monitoring. In most states, the number and frequency of onsite reviews will be limited by available resources. Following a monitoring review, an effective reporting and follow-up process will help ensure that local programs can indeed implement their program improvement plans. Table 4–5 outlines seven steps for effective monitoring, including guidelines and examples for each step.

Table 4–5
Steps and Guidelines for Monitoring Local Programs

Monitoring Steps	Implementation Guidelines	Examples
1. Identify state policy for monitoring. Gather support from those who have a stake in the results.	Provide clear written guidelines to all levels of stakeholders on the scope of the monitoring activities (including process and timelines).	State plan should be open to the public and shared at all levels. State plans often specify: • Outcome measures, and • Frequency of evaluation
2. Specifying the scope of work for monitoring.	Uses quantitative and qualitative data for effective monitoring.	**Quantitative** = look at outcome measurements **Qualitative** = look for evidence using program quality indicators
3. Identify individuals to lead and to participate in monitoring activities.	Consider the unique program features when identifying who should be involved from the local program and who should be part of monitoring team. Consider strength in diversity.	**Local staff:** practitioners, administrators, partners **External team members:** content specialists, other educators, and staff from partnering agencies
4. Identify resources available for monitoring local programs.	With competing demands for resources (staff, time, and money for monitoring), consider formalizing a two-stage monitoring approach.	**Desk reviews** look at program data from a distance. **Onsite reviews** look at data in context—to see first-hand how the process and operations lead to positive outcome measures.
5. Determine process for collecting data with clearly stated criteria for rating. Conduct monitoring activities.	Create and use standard tools for data collection and analysis. Monitors (state staff and team) need to fully understand the tools, their use, and the rating criteria.	**Desk Reviews** can include data, proposals, plans, reports, and program self-review. **Onsite reviews** can include discussion of self-review, observations, interviews, and a review of files and documents.
6. Report on the findings, including recommendations.	Conclude onsite monitoring visits with a verbal debriefing followed by a written report.	Report might include a short description of the monitoring activities with supporting: • Qualitative description • Quantitative information.
7. Follow up on the results.	Given that the major purpose of monitoring is program improvement, is essential and should include an ongoing exchange between the state office and the local program.	Follow-up activities might include reviewing performance standards and program improvement, rewarding or sanctioning, and the beginning of technical assistance.

1. Gather support from those who have a stake in the monitoring results

Effective monitoring requires buy-in among many stakeholders including learners, volunteers, teachers/tutors, counselors, support staff, coordinators, trainers, and local administrators, as well as the institutional leaders and agency partners. To include everyone requires a commitment of time and energy to reach consensus on what really counts and a renewed effort to include multiple measures that capture the richness of programs. When monitoring takes into consideration the interests and concerns of all stakeholders, there is a greater chance for full support during the monitoring process. Everyone needs to understand the process and how data will be collected and analyzed. Stakeholder support will be crucial later for acting on the results or recommendations included in the monitoring report.

2. Specify the scope of work for monitoring

In times of performance-based funding, states need to evaluate local programs to identify the factors that lead to positive results and those that hinder attendance, learning, and progress. Prior to monitoring, everyone needs to fully understand the purpose and expected outcomes. Local programs need to know what the state will review and how they will measure a program's successes and limitations. The monitoring process needs to include a review of qualitative data as well as quantitative data.

3. Identify individuals to be involved

States often use a team approach to local program monitoring so that different levels of staff and partners are involved in the process. When we view monitoring as a capacity-building opportunity, we contribute to statewide professional development efforts and we strengthen our program staff, our networks, and our partnerships. A variety of team members will bring different perspectives to the process.

Local staff are ultimately responsible for the program's success. Local program staff benefit greatly by having a sense of ownership in the process from the planning stages through to the final report. Consider involving the following program staff:

- Administrator, supervisor, or director;
- Coordinators/specialists;
- Staff development team;
- Teachers, aides, and tutors;
- Intake/orientation staff and counselors;
- Data coordinator/data entry staff;
- Support staff;

- Learners (and learner advisory council members);
- Representatives from partner agencies (e.g., employment, social/health services, higher education); and
- Leaders from host agency (e.g., superintendent, president, CEO).

External monitoring teams bring fresh eyes and different perspectives and insights into the process. They might note different strengths and weaknesses of staff/student interactions during observations, they might probe learners and staff more deeply during interviews, and they might bring additional experiences to reviewing data, curricula, lessons, materials, and documents. Involving partners (or staff from collaborating agencies) is especially valuable when a local program serves a significant number of learners with special needs or offers a specific type of program (e.g., employment, workplace ESL, family literacy). Additionally, the inclusion of staff from other programs is a direct way for the state to model peer mentoring and to foster resource exchange. For a minimal cost, monitoring provides a cross-training opportunity and can generate good will.

When the state office is small or understaffed, the state director might designate someone else to lead the monitoring team, for example, a content area leader for reading, math, ESL, special needs, or learning disabilities. Alternatively, there might be a specialist in curriculum, assessment, alternative high schools, or family support. Based on the specific needs or profile of the program being reviewed, a monitoring team might include representatives from the following:

- Other local programs;
- Adult learning resource center or professional development centers;
- Corrections, community colleges and vocational technical programs; and
- Noneducation agencies such as partners from the employment sector or human and social services.

4. Determine the resources available for monitoring local programs

States are continually assessing priorities and resource availability. With the implementation of the NRS accountability measures, states have had an even greater responsibility to manage statewide data and to ensure that local programs are offering effective services, documenting progress, and reporting accurately.

With competing demands and dwindling resources (staff, time, and money for monitoring), many states are formalizing a two-pronged approach with desk reviews and onsite reviews to maximize data gathering while limiting the expenses traditionally associated with onsite visits. Desk reviews allow staff to use information that is regularly collected and reviewed on an ongoing basis without a significant increase in costs or

workload. Costs of expensive onsite reviews costs can be controlled in the following ways:

- Focusing the review on program-specific areas of concern;
- Keeping visit to 1-2 days; and
- Using local, partner, and neighboring staff on the monitoring team.

5. Determine the process for collecting data and rating criteria; conduct monitoring activities

Monitoring activities should focus on both the outcomes and the indicators of program quality. The database can tell us who is being served, but it cannot tell us how they are being recruited or if those who are being served match local demographic needs. For example, a recruitment indicator might state, "program successfully recruits the population in the community identified as needing literacy or ESL services." To collect data demonstrating *successful* recruitment, a desk review could include checking local census data and high school withdrawal rates to determine if the program is reaching non-high school completers. The same sources can provide information related to the need for ESL classes in the community. Onsite interviews with learners, staff, and partner agencies can reveal which recruitment strategies are most effective so that funds can be used judiciously. Monitors might also identify recruitment strategies that are not working for reasons such as dense print, poor translations, inappropriate visuals, or merely because they are not getting into the right hands.

Similarly the database can provide the number and percent of learners who are being pre- and posttested and the rates for those making educational gains. Yet, we also need to collect data on assessment, curriculum, and instruction to identify what strategies are promoting or hindering learner persistence and gain. The criteria specified in the indicators of program quality can help us design interview and observation questions such as the following:

- Are assessment policies being followed?
- How does assessment inform instruction?
- How are learners involved in goal-setting and monitoring their progress?
- Is instruction relevant to individual needs, goals, and learning styles?
- How does instruction promote active learning?

Whether a local program will be part of an onsite review or not, many states ask programs to examine their data continually and to complete a program self-review periodically using the state's program quality indicators. By regularly using both the quantitative and qualitative data, staff can make informed decisions for ongoing operations and continuous improvement.

6. Report on the findings including the recommendations

Anyone who has ever participated in an evaluation knows how stressful it can be both for the program being monitored and for the monitoring team. Keep in mind that the purpose is to improve, so every effort should be made to summarize objectively the most salient patterns that were observed, heard, and noted. An informal debriefing onsite allows the monitoring team and local program staff to discuss the findings and to clarify unknowns or discrepancies before a more formal written report is prepared. Initial recommendations can be shared verbally so that the local program staff can ask questions or request additional assistance.

Most states require a written report for onsite monitoring visits, which can include information from both the desk and onsite reviews. Writers might include a short description of the monitoring activities followed by a qualitative description of findings supported with quantitative information from the data. Anchoring findings to the program quality indicators reinforces the link between program processes and program outcomes. Whenever possible, provide concrete support for the findings by using data tables or charts intermixed with quotes from relevant interviews, citing documented evidence or even the frequency of strengths noted while on site.

Preparing the written review in timely manner is also important. For example, report recommendations will be more useful if they are presented in time for the local program to develop an improvement plan that can be submitted with their proposal. Finally, recommendations for program improvement should be based on objective data, be realistic, and within the capacity of the program.

7. Follow up on the results

Because monitoring is an ongoing process leading to program improvement, follow- up is essential. In many states, local programs are required to respond to the monitoring report with a plan of action within 30, 60, or 90 days. Depending on the report findings and the state model for setting performance standards, the following options might be considered:

- Renegotiating performance standards;
- Reviewing the program improvement plan;
- Providing technical assistance;
- Offering incentives, if possible; and
- Sanctioning, if necessary.

If the primary goal is to foster improvement, make sure the program has the resources and commitment to implement its plan and to foster change. Some programs may need additional assistance in setting realistic objectives and measures. Other programs may need additional technical assistance through professional development, program-to-program coaching, regular consultations, or by directing the program to needed resources such as curriculum, teaching strategies for specific skills, grant writing,

or research on any number of topics (e.g., persistence, language acquisition strategies, performance-based assessments, learning disabilities, or program outreach).

To illustrate many of the key concepts about monitoring presented in this chapter, we conclude with a summary of two states' approaches toward monitoring. Both states use a variety of monitoring methods, including desk and onsite reviews, and they combine NRS data with indicators of program quality to assess program quality. The goal of the monitoring process in both states is to implement a system of continuous local program improvement.

Sample State Monitoring Procedures

Pennsylvania

Pennsylvania's Project Equal is the state's comprehensive project to improve data use among local programs. Program staff review their NRS and local data, identify areas for improvement based on the data, and form a program improvement team. The team develops and implements a program improvement action plan using the indicators of program quality to guide improvements. Staff reviews the data again after implementing the changes.

In addition, Pennsylvania uses its indicators of program quality as a framework for local program evaluation and self-assessment. They guide the bureau's program monitoring during onsite reviews. The indicators are used as a tool to prioritize and guide continuous improvement. Using data and the indicators of program quality helps program staff gain valuable insight into their services and outcomes.

1. **Example: Leadership and Continuous Improvement**—Indicator 3.1: Program leaders ensure full staff participation in the continuous improvement process.

   Monitoring through Desk Review: *Document Review of Program Improvement Plans*

   - Do the data indicate improvements in areas designated in the Program Improvement Plan?

   Onsite Review: *Interviews*

   - Are leaders and practitioners aware of the continuous improvement process (CIP)?
   - Do they know who is on the CIP team? Is staff paid to attend CIP team meetings?

2. **Example: Instructional System—**Indicator 2.6: Curriculum provides for learning and skill building in real-life contexts that allow learners to become effective family members, workers, and citizens.

   On Site Review: *Observations, Interviews, and Document Reviews*

   - How is the content of curriculum and lessons determined? Is there a process for initiating curriculum and content changes?
   - Is curriculum available? Does curriculum include real-life contexts?

- Do lessons incorporate real-life contexts and materials?
- What books and resources are available for staff? Are they linked to curriculum?
- Do instructors have records of what is taught?
- Do instructors document learner progress?

Tennessee

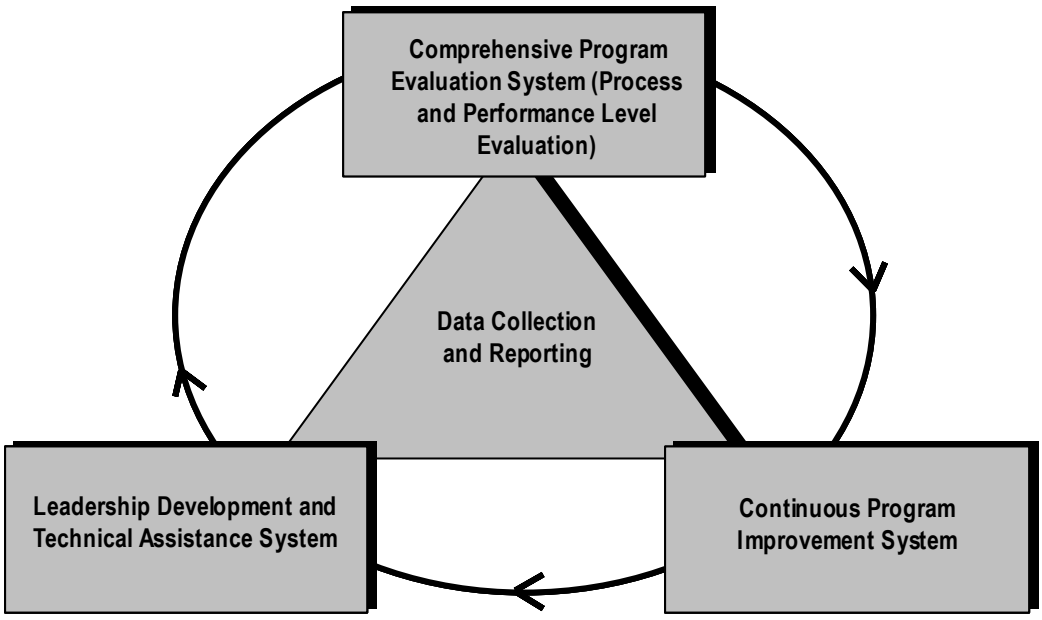
Strategy for Program Evaluation and Continuous Program Improvement

In the 1999–2000 program year, the Office of Adult Education adopted a quality/continuous improvement model called "Quality AE: Tennessee's Continuous Program Improvement Initiative." As part of the model, programs will be evaluated through the following steps:

1. Collect local provider and student performance measures;
2. Determine the level of student performance improvement;
3. Identify program quality; and
4. Determine the extent to which populations specified in the state plan were served.

The initiative integrates program reporting and data collection, program evaluation, continuous improvement, and leadership development and technical assistance into a comprehensive model. Exhibit 4–1 outlines Tennessee's overall strategy for program evaluation and continuous program improvement. The graphic below shows each component of this model.

Exhibit 4–1
Overall Strategy for Program Evaluation and Continuous Program Improvement



The comprehensive program evaluation system integrates different types of program evaluation information. This includes (1) a process level evaluation of the activities and services provided by local contractors, and (2) a performance level evaluation based on statistical performance data addressing learner and program outputs, outcomes, and impacts.

1. **Measures of Program Quality:** Program Quality Indicators (PQI) evaluation instrument provides a way of assessing local programs at a process level—an assessment of the activities and services that local programs offer. The Office of Adult Education staff will annually monitor the self-review process and verify results. This review will include targeted site visits and technical support to local providers as needed. Local programs will develop a program improvement plan annually based on the results of self-assessment.

2. **Statistical Performance Data:** All local providers will be required to complete comprehensive statistical performance reports at the end of each program year. These statistical performance reports will also address the core and secondary performance indicators and measures, which include improvement in skills levels, placement or retention in employment, and receipt of a diploma as well as other secondary measures. The statistical performance report will also track attainment of learner-articulated goals, customer satisfaction levels, and computer literacy measures.

3. **Student Follow-up Studies:** The Office of Adult Education will identify strategies that local providers can implement to follow up on students who leave the program before completing their goals, as well as for students who leave the program after meeting their goals.

4. **Target Population Penetration Studies:** The Office of Adult Education will annually estimate the cumulative market penetration of the Adult Education and Family Literacy Program on salient populations targeted in this state plan.

Chapter 5. A Model for Program Improvement

We have emphasized in this guide that the ultimate goal of accountability is program improvement. Accountability systems affect program processes through the setting of performance standards, the standard-setting model and the system of rewards and sanctions that the state implements. When accountability systems are well implemented, the performance data provide accurate indicators of performance and of the program processes under the data. Through desk reviews of data and onsite monitoring, the state can identify which program elements are working well and which need improvement.

This chapter addresses the final step of the data driven program improvement process (see table 5–1), designing and implementing an improvement plan based on what you have learned through program monitoring. Two principles guide our discussion of program improvement. To be effective and meaningful, program improvement should:

- Be a continuous activity built into the accountability process, not a one-time or episodic event; and

- Be a multidimensional process that varies by approach, implementation, and outcome.

First we present a rubric of continuous program improvement that illustrates the program improvement continuum, and then we present a program change model to guide the planning and implementation of program improvement. Concluding this chapter is an example of how one state is using NRS data as a guide for implementing a local program improvement plan.

The Program Improvement Process

Program improvement requires real change in how things are done. Regardless of what program elements need changing or where your state is on the program improvement continuum, you must initiate a process of change if you want to see a difference in performance. This process is never easy, but is possible when it is well planned, sets realistic goals, offers sufficient resources, and clearly defines the anticipated outcomes of the change process.

A Continuum of Program Improvement

Table 5–1 is a rubric for measuring continuous program improvement adapted from *Continuous Improvement Continuum*, which is based on the Malcolm Baldrige Award Program. The rubric serves both as a starting point for program improvement by focusing discussion on where a program stands along the continuum of program improvement and as a goal to work toward.

Table 5–1
Continuous Program Improvement Rubric

Program Improvement Continuum →

	One (absent)	**Two**	**Three** (Integrated)
Approach	No quality plan or process exists for program improvement. Data are neither used nor considered important in the program improvement process.	A comprehensive program improvement plan is in place. Plan includes action steps that move the program towards improving student outcomes.	The plan for continuous program improvement has a clear focus on student outcomes. There is a clear articulation and integration of all program policies and procedures. The program improvement team ensures that all action steps are implemented by the appropriate parties.
Implementation	There is no knowledge of or direction for quality program improvement.	Implementation goals, responsibilities, and timelines are spelled out. Supporting structures for implementing program improvement exist.	Program goals and performance standards are shared and articulated throughout the program. Planning and effective instruction that meet students' needs are linked to performance standards. Evaluation of improvement efforts is an important part of the process.
Outcome	The response to failures and mistakes is finger pointing and blaming others. Mistakes are repeated. No coordinated effort exists to address program deficits.	There is evidence that the program improvement plan is being implemented in some areas of the program. Improvement is not yet systematic or fully integrated programwide.	Significant improvement in student learning is a result of quality program improvement for all aspects of the program. The staff understands and shares the importance of quality program improvement and accountability.

Adapted from Data Analysis for Comprehensive Schoolwide Improvement (1998)

The rows represent three stages of the improvement process: approach, implementation, and outcome. Each of the columns represents an example of what a program at a particular point on the program improvement continuum might look like. So, for instance, a program at stage *two* would exhibit the characteristics outlined in the rubric for that stage in its approach to continuous program improvement, its implementation of that approach, and in the outcomes to implementing that approach. A *one* represents an absence of a program improvement process, while a *three* represents a highly developed program improvement process. The goal for a program is to move toward stage *three* in its program improvement. The key questions to focus on are as follows:

- Is quality data used in the program improvement process? If not, what needs to happen to get quality data?
- Is there a focus on student outcomes? If so, how is the outcome focus demonstrated and documented?

- Is there alignment and coordination of program polices and procedures toward the end goal of improving performance? If not, what areas are most out of alignment and in most need of immediate attention?

A program that is at stage *three* focuses on performance standards that measure student outcomes with quality data, coordinates different aspects of the program, and exhibits clear communication among staff to reach those performance standards. Finally, in such a program, program improvement is a continuous process that includes an evaluation of the outcomes resulting from the program improvement process.

A Model of the Program Improvement Process

While the Continuous Program Improvement Rubric provides a way of assessing your program on a Program Improvement Continuum, Exhibit 5–1 shows a model for the steps to take to move your program along that continuum. It divides the program improvement process into four steps:

1. Planning;
2. Implementing;
3. Evaluating, and
4. Lessons Learned.

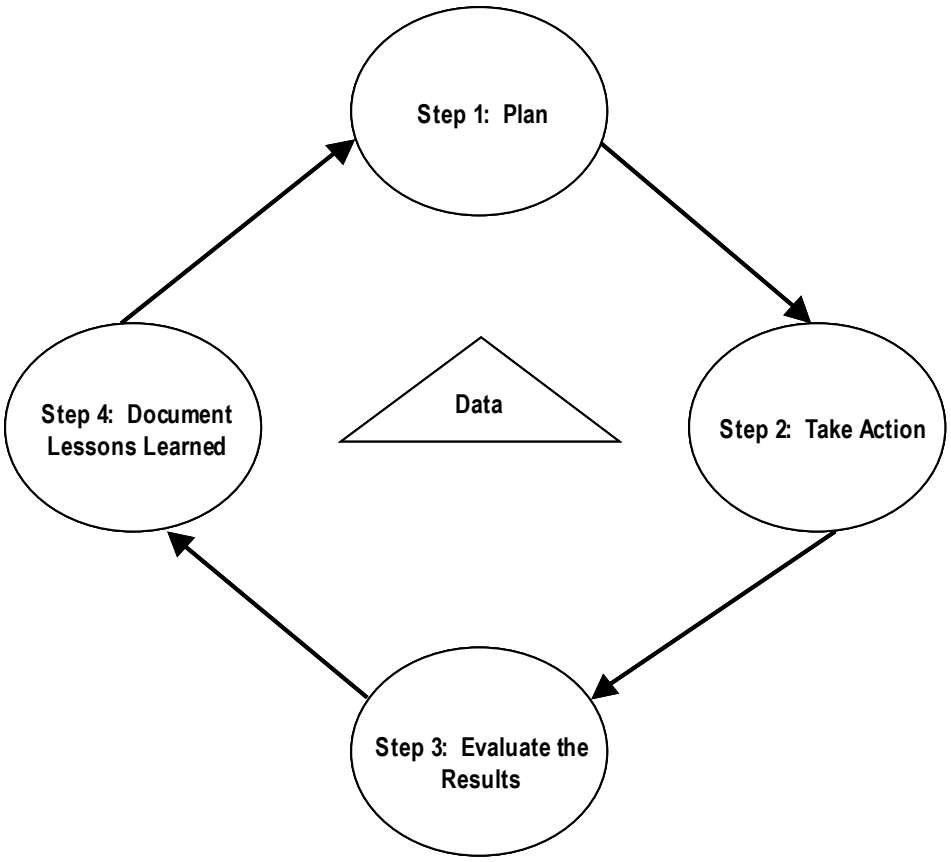
At the center of this process are data–both data collected during the monitoring phase discussed in chapter 4 and the new data collected during the program improvement phase. Although the exhibit shows the program improvement process as progressing from one step to the next in an orderly fashion, the reality is often much messier, with arrows going in both directions and the steps often out of sequence. For instance, after implementing your strategy, you may encounter unforeseen problems and return to planning. Recognizing the reality, we present an idealized discussion of each step.

Plan Your Strategy

Develop a Program Improvement Team

Similar to the development of a monitoring team, you will need to develop a team for program improvement. Table 5-2 indicates some of the important people to consider for your program improvement team and what they bring to the process. Individuals may play more than one role and bring a perspective other than that usually associated with their position. For instance, the program director may have a good sense of what goes on in the classroom day to day and a teacher may have a good sense of the big picture.

Exhibit 5–1
The Program Improvement Process



The program improvement team serves several functions. First, it brings together a group of people with various perspectives and expertise to help design and implement a program improvement plan. Second, it helps identify and acquire resources for implementation of that plan, and finally, it ensures buy-in during and after implementation. This last function is especially critical for program staff. Involving staff from the beginning will increase ownership and an understanding of what it takes for implementation to succeed. Also, staff on the front line possess knowledge of the day-to-day workings of the program that is invaluable for the team to understand the story behind the numbers.

Prioritize Areas Needing Improvement

The next task for the program improvement team is to prioritize the areas to be addressed. Likely, the team will identify more than one area needing improvement. Given resource limitations, however, the team has to identify a reasonable number of areas to focus on, and of those, which areas should be addressed first. Setting clear priorities allows the team to focus its resources on those areas leading to the greatest program improvement. Good data will aid you in targeting limited resources (both time and money) so that they have the biggest impact.

Table 5–2
Program Improvement Team

Staff	Contribution
Program Staff	
Supervisor/Director	Ultimate responsibility for implementation.
Teachers	The needs of the students. Content expertise. Responsible for direct implementation.
Professional Development Staff	Curriculum and professional development expertise. Likely to have role in implementation.
Student support and administrative staff.	Issues the students face that affect their learning. Administrative functioning of the program. Likely to play a role in implementation.
Learners	The needs of those being served.
External Partners	
Directors of other programs	Inside knowledge of another program. Possible knowledge or access to sources of funding.
Relevant staff from other programs	Inside knowledge of another program.
Governmental Agencies	Expertise in program improvement process and content areas. Possible source of funding.
Nongovernmental Agencies/Foundations	Expertise in program improvement process or content areas. Possible source of funding.
University Faculty/Researchers	Expertise in content and program improvement process. Possible knowledge or access to sources of funding.

How does one go about prioritizing? The probability that the identified program areas will promote program success with the available resources will help you make the decision. In addition, the idea of alignment embodied in the Program Improvement Continuum may be helpful here. The objective is to align all elements of the program so they are working toward the same end, rather than at cross-purposes. The program goals set the program direction, and the various program elements should move the program toward those goals. The data gathered in monitoring will help identify which elements are out of alignment. Asking the following questions may help you prioritize areas to change:

- Is improving this area critical to the success of the program?
- What will it take to improve in this area?
- What are some of the obstacles that stand in the way of improvement in this area?
- Are there other areas that need to be addressed before this area can be addressed?
- What resources will it take to implement a particular strategy? What is the benefit relative to the cost in resources (remember resources include time and the goodwill of the staff in addition to money)? How does this compare to alternative strategies?
- Do we have these resources? If not, what is the chance of acquiring them? How will you go about getting resources? One advantage of involving others from

outside the program early on is that they may be in a position to help you acquire resources.

The primary objective when you are prioritizing is to approach program improvement as a strategic process, rather than merely responding to events. The box below offers an example of setting priorities for improvement.

Gathering and Selecting Promising Strategies

After prioritizing, the team should identify the appropriate strategies with which to address areas identified as needing improvement. In its search, the team should cast a wide net. Examine numerous sources to gather ideas for appropriate strategies. These sources include:

- Other programs
- Peers
- Research
- Experts in the field

Other programs and peers. Identify other programs with a population similar to yours that have successfully addressed problems similar to the ones you face. Although no two programs are identical in the populations they serve and the goals they set, it is still instructive and useful to look around and see what other programs are doing and how their approach can help you improve your own program. Talk to your peers in other programs to see what has and has not worked for them. The more you can build on the success of other programs and learn from their mistakes the more likely your improvement efforts will succeed.

Research and experts in the field. With the current emphasis on using research-based practices to guide program improvement, it is important to examine the research literature for promising practices. Identify and use the appropriate experts as guides in navigating the research. It is not easy to transfer research to practice, so use a critical eye when examining the quality and appropriateness of possible strategies to use in your own program. Although there are no hard and fast rules for judging the quality of various strategies, below are some questions that may help you make a more informed judgment:

- How appropriate is the strategy to your program?
  - Has it been used with a similar population and for similar purposes as your intended use? If not, can it be modified to serve your programs needs?
  - If the strategy were successful, would it advance you towards your program goals? If all the strategy can do is move you more quickly in the wrong direction, discard it.
- How effective is the strategy?

- Has the strategy proven successful with other programs over time? Is it clear that the particular strategy is responsible for the success and not some other aspect of the program?

- Is there research on the strategy? Is the research from a reliable source? Has it been shown to work over time and in different situations? Was a controlled experiment used to judge the effectiveness of the program?

- If the strategy is new, is there research indicating certain components of the strategy are effective? Is there anything in the research literature about the program's probable effectiveness?

The more information you gather on the above questions the better. Seldom, however, are the answers clear cut or simple, especially in something as complex as adult education. You will need to use your best judgment. The more that judgment is informed by the use of quality data, the more likely you are to make decisions that improve your program.

Taking Action

Implementing Change

Once you have finished prioritizing the changes to be made, examined the available resources, and determined how to get them, it is time to implement the changes. Although staff buy-in is important during the whole program improvement process, it is especially critical here. It is at the point of implementation that the staff begins to feel the real impact of the changes. You may be asking people to change behaviors and procedures that they have become comfortable with and that have worked for them. How much you should involve different staff depends on the extent of the change and the degree of their involvement in the change. For instance, if the change involves a particular part of the curriculum, those most directly involved will be those who teach that curriculum. Because different parts of a program are intertwined, it is important to consider all those who are affected. There is a balance to be struck, though, between involving staff in changing something that is important to them and overburdening staff with matters they feel have no impact on them. Staff's time is a resource to use wisely.

Leadership. Strong and effective leadership is important for establishing a clear vision, bringing staff on board, and marshalling the necessary resources. Clearly communicating the big picture, the reasons for the change, and what the program will look like once change is implemented is essential to successful buy-in and improvement. A good leader helps focus people on the task at hand and provides a vision that motivates them to accept the disruption inherent in change. No matter how effective the leadership, however, people will differ in their responses to change. Some may adopt and adapt more readily than others. Use these people as advocates for the change.

Promoting and nurturing the change. Change is not a single event, rather an ongoing process. Successful change requires continuous attention and tweaking as the process unfolds. It is important to remember that the progression of change is usually not a neat, orderly event.

Consequently, you need to monitor the implementation process. This oversight provides the improvement team with a sense of how implementation is unfolding and what needs to be modified. Use data, observation, and communication with other staff to get a sense of what is going well and what is not. Clear and continuous communication and feedback from the staff are critical elements to the whole change process. Think of communication and feedback as another form of data. Like all data, it is useful only if it is forthcoming and accurate. Creating an environment in which critical feedback is truly valued will increase the flow of accurate information. The purpose of monitoring the change process is not to judge or evaluate people, but to monitor the change process itself.

Evaluate the Results

The evaluation of the outcome of program improvement activities is often overlooked. This activity focuses on whether program outcomes have changed as a result of the new procedures and if there have been any unintended consequences. Below we list some general considerations to keep in mind when you are designing an evaluation and interpreting the results.

Consider both short-term and long-term evaluations. The more complex the changes are, the longer it will take for the impact to show and be measurable. How does one determine when to evaluate and for how long? The answer depends on a number of factors, including the nature of the changes made. Changes that seek to alter numerous elements of the program in substantial ways can take years to take hold. Even seemingly simple changes may take awhile to show an impact because education is a complex enterprise that involves changing people's behavior.

For example, making changes in instruction is really a set of interrelated changes, and its success depends on the interplay of all those changes. Both teachers and students need time to become comfortable with the new approaches and adapt it to their needs. Teachers not only have to understand what is new, but they must also effectively communicate it to their students. These changes also will require ongoing training for teachers so they can adapt the new approaches to their own teaching style.

Another reason to consider both long- and short-term evaluations is that short-term changes may not always accurately reflect long-term trends. Change can cause short-term disruption that may result in an initial negative impact, but would have a long-term positive impact if left in place long enough to show the positive results. Initial negative evaluation results may reflect a new learning curve as staff and students adapt to the changes. We have all experienced the frustration of trying to learn a new way of doing something. There is usually a period in which we actually do worse than we did before the change. On the other hand, a short-term gain may reflect initial excitement and focus brought about by changes that, with time, may dissipate.

Interpret the results of the evaluation carefully. Ask yourself whether the improvements you observe really are the results of changes you made or are the results of some other, perhaps unrelated, changes. For instance, if while modifying instructional approaches, you also recruited more motivated students, the improvement you observe may reflect the change in

student motivation, rather than the change in instruction. Even if you identify the changes you made as the cause of the improvement, it may still be difficult to identify which part of the strategy is responsible for the improvement, especially if the strategy has many component parts.

Look for unintended consequences. Often unintended consequences sprout up when you are making changes. For instance, changing the curriculum may lead to some teacher dissatisfaction. Because different parts of a program are so interconnected it is often hard to predict or limit the changes that will occur. Sometimes, you are lucky and the unintended consequences are positive, but often they are not.

After analyzing the results of the evaluation, the team will need to decide whether the changes have resulted in program improvement and whether to stay the course. Depending upon the results of your evaluation, you have several options:

- ***Significant improvement with no significant negative unintended consequences:*** Stay the course. The changes are having the impact you were hoping for. You will want to evaluate at a later point to determine if this positive trend continues.

- ***Little or no improvement:*** Consider if you have allowed sufficient time for improvement to take place. You may want to stay the course and reevaluate within a reasonable amount of time to see if improvements have started to take hold. Keep in mind that in education, reasonable strategies are often not given a reasonable amount of time to have an impact. Discarding a possibly promising strategy can be as harmful as retaining an ineffective one. It needlessly wastes resources, including the staff's willingness to change in the future. If you think there has been significant time for improvement and none has occurred, think of alternative strategies for improvement. Consider whether the problem lies more with implementation or the strategy itself.

- ***A deterioration in outcomes:*** The most obvious choice is to scrap the changes and think of alternatives. Depending on the size of the negative impact, consider whether this is a temporary setback as people adjust to a new way of doing things, or a real deterioration.

Documenting the Process: Lessons Learned

Documenting what you have done and learned during the program improvement process is an important step. Unfortunately, programs often skip this step, due to resource limitations and competing demands to address more imminent needs. However, while documenting what you have done may feel like an unnecessary burden, it is more than an academic exercise. It will save valuable time when you inevitably go through the process again. Documenting the steps you took, what worked and what didn't, and what you learned in the process will serve as a guide for the future. You should think of this as an evaluation of the program improvement process itself. Although this step is placed at the end of the process, documentation and evaluation should be done from the beginning through to the end. If you document what you have done as you go, you will retain important details that would otherwise fade over time.

Documenting the program improvement process also will allow you to share lessons learned with other programs. Not only will you help them in their program improvement, but also you will help establish a culture of sharing information and strategies. Program improvement is a continuous process that requires constant attention. The more you pay attention to the process and learn from your successes and mistakes, the more likely it is that your program will improve.

The use of data in the program improvement process and some of the key concepts presented in this chapter, are illustrated in the following box. This example shows how one state approaches program improvement by providing the local programs with NRS data. The state guides them with a series of questions to consider as they review their data and outline their program improvement process.

Massachusetts: Program Improvement Strategy

The Massachusetts Department of Education (MDE) provides customized reports for adult education programs to assist and guide them in their program improvement process. The reports include NRS measures of student attendance, educational gains and goal attainment and the related performance standards. A series of guide questions accompany each set of data. The guide questions are designed for programs to use in the program improvement process. To use the data effectively for program improvement, the MDE suggests the following steps:

Start With the Data

- Look at the data in this report;
- Consider the guiding questions;
- Create additional questions; and
- Identify other relevant data.

Create a Plan

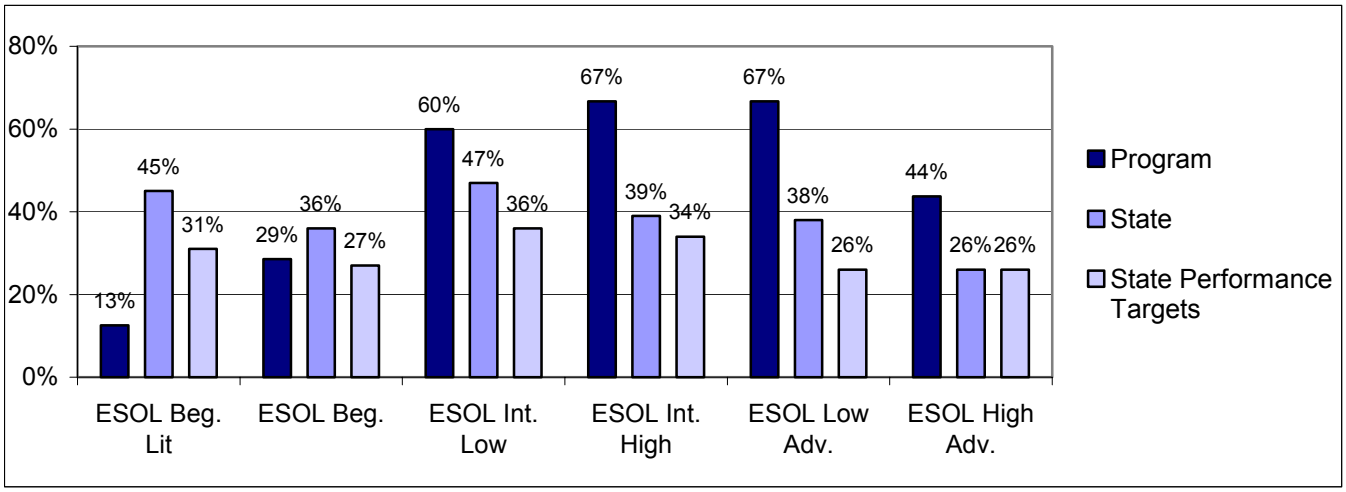
- Prioritize areas for improvement;
- Research successful strategies;
- Make a plan;
- Implement it; and
- Evaluate what you have done.

The customized report shown below contains a sample of data regarding completion rates for the ESOL levels. Exhibit 5–2 and presents data on the number and percentage of students completing each level within the program and for the state as a whole, as well as the state performance targets. The two different formats are useful in highlighting different aspects of the data. The bar graph quickly communicates the relative standing of the program in relation to the state average and the state performance targets. The table provides more detail, displaying the underlying numbers in addition to the percentages, but does not as readily lend itself to comparing the program with the state average and state targets.

The bar chart suggests that the program is performing well in the intermediate ESOL and advanced ESOL levels, but poorly at the beginning ESOL and beginning literacy ESOL levels in relation to other programs and other levels. The table provides additional information such as the number of students underlying the percentages. The table allows you to see, for instance, that only a total of three students are enrolled in the low-advanced level and seven students are enrolled at the beginning level. With so few students at those levels (and others), the success or failure to complete a level of one student alters the percentages dramatically. If just one more student had advanced out of the beginning level, the percentages for level completion would have shot up from 29 percent to 43 percent.

The guiding questions encourage the program to generate explanations that identify program processes and data collection procedures under the data and to seek additional data to evaluate the plausibility of explanations. This process helps the program to identify areas of concern and to develop a program improvement plan to correct problems and improve performance.

Exhibit 5–2
Sample Data Report
Educational Gain—ESOL Levels



NRS Level	Number Completing Level and Total Enrolled				% Completing Level		State Performance Targets
	Program Completed	Program Total	State Completed	State Total	Program	State	
ESOL Beg. Lit.	2	16	1378	3,057	13%	45%	31%
ESOL Beg.	2	7	1363	3,770	29%	36%	27%
ESOL Int. Low	6	10	936	1,992	60%	47%	36%
ESOL Int. High	6	9	743	1,903	67%	39%	34%
ESOL Low Adv.	2	3	623	1,626	67%	38%	26%
ESOL High Adv.	7	16	418	1,590	44%	26%	26%
TOTAL	25	61	5,461	13,938	41%	39%	

TOTAL ABE/ESOL	**159**	**270**	**7,931**	**24,488**	**56%**	**32%**

Guiding Questions

- What is your goal for educational attainment for each level?
- Where does your program perform well? What can you learn from that?
- What factors contribute to your students' educational gain?
- What additional information do you need (e.g. curriculum, attendance, instruction, staffing, student educational background)?

Bibliography

Alamprese, J. and Stickney, E. (1999). *Development: Lessons Learned from Project Educational Quality in Adult Literacy (EQuAL).* Cambridge, MA: Abt. Associates.

Annenberg Institute for School Reform at Brown University. (1998). *Using Data for School Improvement: Report on the Second Practitioners' Conference for Annenberg Challenge Sites*. Available at: http://www.annenberginstitute.org/images/using_data4.pdf.

Baldrige National Quality Program. (2004). *Education Criteria for Program Excellence.* Gaithersburg, MD: National Institute of Standards and Technology.

Bernhardt, V.L. (1998). *Data Analysis for Comprehensive Schoolwide Improvement*. Larchmont, NY: Eye on Education.

Bingman, M. and Ebert, O. (2001). *How are we Doing? An Inquiry Guide for Adult Education Programs.* NCSALL.

Bond, S., Boyd, S., Rapp, K. (1997). Taking Stock: A Practical Guide to Evaluating Your Own Programs. Chapel Hill, NC: Horizon Research, Inc.

Brooks, S. (2000). *How States Can Hold Schools Accountable: The Strong Schools Model of Standards-Based Reform*. Center on Reinventing Public Education, Daniel J. Evans School of Public Affairs, Seattle: University of Washington.

Caffarella, R. (1994). *Planning Programs for Adult Learners, A Practical Guide for Educators, Trainers and Staff Developers.* San Francisco: Jossey-Bass Publishers.

Cody, J., Ford, J., Hayward, K. (1998). A Story of Improvement. *Focus on Basics,* Vol 2, Issue C. Cambridge, MA: NCSALL.

Comings, J., Beder, H., Bingman, B., Reder, S., Smith, C., (2003). *Establishing an Evidence-based Adult Education System.* Cambridge, MA: NCSALL.

Elmore, R. (1997). Accountability in Local School Districts: Learning To Do the Right Things. *Advances in Educational Administration, 5,* 59-82.

McLendon, L. and Polis, K. (2001). *The Adult Education State Director's Going to Scale Guide for Planning, Improving, and Evaluating State Program Improvement Initiatives.* Washington, DC: National Adult Education Professional Development Consortium.

Merrifield, J. (1999). Performance Accountability for What, To Whom, And How? *Focus On Basics,* Volume 3, Issue B. Cambridge, MA: NCSALL.

Olson, L. (2004). ETS Imports 'Formative Assessment' Analyst. *Education Week, 23*, No. 4, February 25.

Simpson, E. L. (1987). An Interactive Model of Program Development. In C. Klevins (ed.). *Materials and Methods In Adult and Continuing Education.* Los Angeles: Klevens Publications, Inc.

Stecher, B. (1995). *Improving Performance Measures and Standards for Workforce Education.* Berkeley, CA: National Center for Research in Vocational Education.

Trott, C. and Baj, J. (1996). *Building State Systems Based on Performance: The Workforce Development Experience. A Guide for States*. Washington, DC: National Governor's Association.

Web Sites

Community Partnerships for Adult Learning: Program Evaluation
http://www.c-pal.net/build/assess/how.html

Kansas
http://www.kansasregents.org/download/adultEd/indicators.pdf

Ohio
http://www.ode.state.oh.us/ctae/ABLE

Oregon
http://www.ccwd.state.or.us/abs/default.htm

Pennsylvania – Indicators of Program Quality Updated 1999-2000
http://www.able.state.pa.us/able/lib/able/ipq.pdf

Tennessee
http://aeonline.coe.utk.pqi.htm

Appendix

Indicators of Program Quality and Self-Review Instruments

In the early 1990s the National Literacy Act (NLA) required states to develop model indicators of program quality for adult education. Using the model standards developed by the Office of Vocational and Adult Education, most states developed quality indicators to define good program policy and practice. Many states continue to use these indicators to guide their program monitoring and program improvement efforts. In the context of NRS outcomes and their relationship to program processes, quality indicators are useful for evaluating program quality and for tying the standards into the accountability systems. Indeed, several states already do this as part of their approach to program monitoring.

As defined by the NLA, quality indicators measure "efficient and effective performance" of adult education programs, so they are extremely valuable as part of our monitoring toolkit. The following pages provide short clips from several state program self-review and document instruments. They illustrate a variety of formats and criteria that have been developed and refined over time. You will note that some states rely exclusively on quantitative measures while others combine both qualitative and quantitative standards.

Whenever, possible we have included the Web site link to the full document for your review and consideration. We hope these samples will stimulate a critical review and make the process easier for states that will be updating or revising their indicators for monitoring purposes.

Tennessee provides suggestions for documents that can be reviewed along with specific criteria for rating each measure.

http://aeonline.coe.utk.pqi.htm

Indicator 1. *Learners demonstrate progress toward attainment of basic skills and competencies that support their educational need.*

Measure	Documents	Points	Comments
A. Pre- and retest scores on standardized tests are used as a way to demonstrate progress.	• ABLE • TABE • AMES • Practice GED • GED • BEST (ESOL) • CASAS (ESOL)	0. No testing documentation 1. Pretests on some students but not all 2. Pretests on all enrolled students 3. Pretests on all enrolled students and retests on up to 25% of students (tested at appropriate intervals or on completion of program) 4. Pretests on all enrolled students and retests on more than 25% and less than 50% of students 5. Pretests on all enrolled students and retests on 50% or more of students	
D. Students are involved in their own goal-setting and assessment of progress.	• Student files • Course critiques • Self-completed progress checklists • Notes on student contact sheets • Relevant questions on intake forms • Student biographies or essays	Student goal-setting: 0. No documentation. 1. Some documentation. 2. 25% of students have evidence of goal-setting/progress. 3. 50% of students have evidence of goal-setting/progress. 4. 75% of students set goals which are used in designing program, goals reassessed and restated at intervals, documentation of periodic assessment.	

Kansas has developed a simple point system for rating each measure (see top measure) and also has a checklist for document reviews as part of the monitoring process. http://www.kansasregents.org/download/adultEd/indicators.pdf

2. The program is of sufficient intensity and duration so that participants *demonstrate progress* toward their educational goals [AEFLA, Section 231, (e)(4)(A) and (e)(7)].

Measure 2.2: Participants make significant educational gains, confirmed through nationally standardized assessments.		
	High (4 points)	70% or more of all participants who complete an education level demonstrate completion through CASAS pre- and posttesting.
	Medium (2 points)	40% to 69% of all participants who complete an education level demonstrate completion through CASAS pre- and posttesting.
	Low (0 points)	Less than 40% of all participants who complete an educational level demonstrate the completion through CASAS pre- and posttesting.

CASAS pre- and posttesting: Administration of a CASAS Reading, Math, or Listening Diagnostic Test or CASAS Functional Writing Assessment as a pretest followed after an appropriate amount of instruction by an alternative version of the same assessment. Note: This measure has been significantly revised.

Program Documentation

Programs should have the following documents, relative to the specified Indicator and Measure (I:M), on file or readily accessible. For each type of document, √ whether it exists (yes, no, or partially), and add any remarks as applicable.

Doc #	I:M	Document	Yes	Part.	No	Remarks
2	1.2 1.3 4.2	Evidence of **recruitment** of low-income adults, adults with minimal literacy skills, single parents, adults with multiple barriers, criminal offenders, institutionalized individuals				
4	1.2	CASAS Appraisal/Diagnostic Tests to determine program entry classification levels and to aid in developing learning plans				
5	3.1	Learner attendance records reflecting hours reported on ABE2003				

Oregon requires the reviewer to note specific evidence and then rate each indicator according to state performance standards.

http://www.ccwd.state.or.us/abs/default.htm

8. Instruction maximizes learner and program attainment of goals.

Strategy II. Identify and define content of instructional programs

Process	Indicator	Evidence	Exceeds	Meets	NI	NO
Program identifies and defines the content of each instructional program.	1. Program has curriculum on file that includes the following descriptions: • Instruction purpose; • What learners will know and will be able to do; and • How learner outcomes are assessed					
	5. Instructor and students develop individual education plans that include the following: • Learner goals; • Assessment results at entry; • Ongoing assessment outcomes/gains; and • Periodic reviews and updates to document learner progress.					

NI – needs improvement
N0 – not observed

Pennsylvania uses a more open-ended review instrument that leads directly into planning program improvement.

http://www.able.state.pa.us/able/lib/able/ipq.pdf

1. Customer Results and Program Accountability

Key Concept: Quality is ultimately judged by learner outcomes and customer satisfaction.

1.2 Learners remain in the program long enough to meet their educational goals

I. Discussion Questions for Program Improvement Teams

1. How are learner goals determined and documented within agency? Who is responsible?

2. Where is learner goal information maintained?

3. How often is learner goal information updated? Who does the updating?

4. How is goal attainment tracked within the program? How do we make sure it gets entered in Lit Pro?

5. What is our agency's student goal attainment rate?

II. How Is our Agency Doing With This Indicator:

** Agency systematically collects learner educational goals and attainment that is recorded, reviewed, updated and tracked. Goal attainment is determined from the data collected.*

Evidence: _____

_____

_____

_____

Based on this evidence, our agency would score as follows on this component:
3 Strong evidence 2 Some evidence 1 Minimal evidence 0 No evidence

III. Possible Data for Decision-Making Questions Related to This Area

- How can we help our students better articulate short and long-term goals?
- What goals do our learners most often pursue?
- How can we improve our agency goal attainment rate?

IV. Next Steps for our Agency Related to Indicator 1.2

_____

_____

_____

Maryland's Program Self-Review instrument is not available on the state Web site, but we have included a sample since Maryland used an electronic survey, which allowed a variety of local program staff members to participate anonymously. A major advantage for using an online survey is that the results are tabulated and reported electronically. The sample below includes the results from 11 staff members in one rural county.
For a free sample, visit www.zoomerang.com.

Assessment Standard 2: Multiple assessment tools and strategies are used to gather information and to guide program design and curriculum development

Indicators	NA	Not evident	Beginning	Developing	Accomplished	Exemplary
1. Learners participate in the identification of the skills they need, and the articulation of their learning goals (e.g., life skills or skills for parenting, for the workplace, for using technology, etc.).	9% (n=1)	0% (0)	0% (0)	0% (0)	55% (6)	36% (4)
5. Locations/schedules for educational programs are determined through a survey of the target population.	27% (n=3)	18% (2)	0% (0)	27% (3)	9% (1)	18% (2)