PART SEVEN

Evaluate Measurement

This part of the Guide describes several approaches for evaluating the components of a measurement program to ensure that they are being performed as intended and to identify potential improvements. This part explains how to evaluate the measures and indicators, as well as the measurement process. It consists of five chapters:

- **Chapter 1, *Evaluate Measurement* Overview,** provides an overview of this activity's four steps.
- **Chapter 2, *Evaluate Measures and Indicators*,** provides criteria for evaluating the products of the measurement process.
- **Chapter 3, *Evaluate the Measurement Process*,** describes three approaches to evaluating the measurement process.
- **Chapter 4, *Update the Experience Base*,** introduces the experience base as a repository for lessons learned and reusable measurement assets.
- **Chapter 5, *Identify and Implement Improvements*,** describes how improvements are selected and how their implementation is coordinated with the other measurement activities.

[This page intentionally left blank.]

1

Evaluate Measurement Overview

The objective of the measurement program is to generate information that provides insight into project issues so that the project manager can make informed decisions. It is unlikely that the first implementation of a measurement program will accomplish this perfectly; experience will uncover better measures and methods of processing and analyzing data. Evaluate the measurement program regularly to continually improve it.

The *Evaluate Measurement* activity includes four tasks, as shown in Figure 7-1. The first task, *Evaluate the Measures and Indicators*, considers whether the measurement data and indicators satisfy the project manager's information needs. The second task, *Evaluate the Measurement Process*, examines the efficacy of the project's measurement process. The evaluation of a measurement process involves three dimensions:

- Quantitative performance of the process
- Conformance of the measurement process to the measurement plan
- Capability (maturity) of the measurement process relative to a standard

The third task, *Update the Experience Base*, identifies lessons learned from evaluating the measurement products and process. Some of these lessons may be implemented on the current project, while others may benefit future projects. The final task, *Identify and Implement Improvements*, identifies specific actions to improve the project's measures and indicators, or the measurement process. Improvements may be made by updating the measurement plan in the next cycle through the *Tailor Measures* activity, or as a parallel improvement initiative.

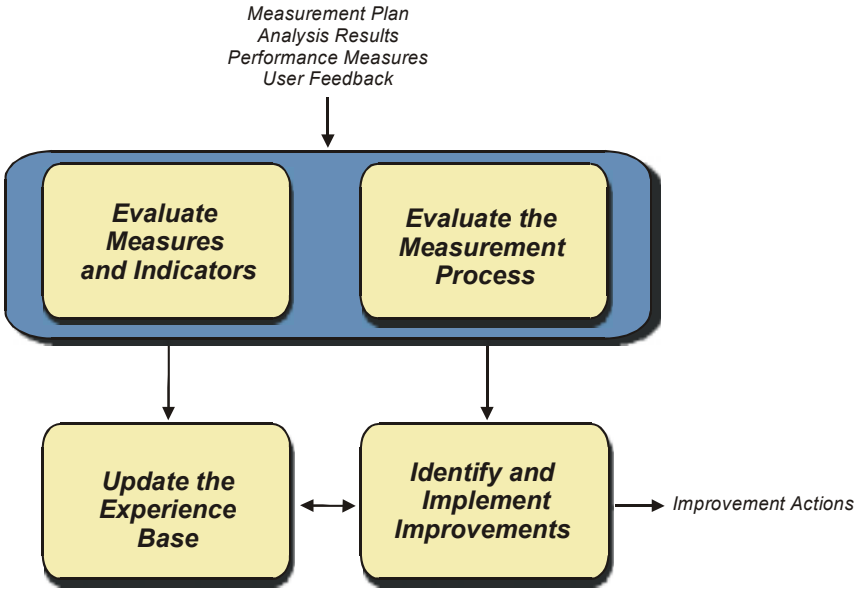

Figure 7-1. *Evaluate Measurement* Activity

Evaluate the measures and indicators as well as the measurement process regularly. Planning these evaluations in advance helps to ensure that they occur. Evaluation plans may be part of the measurement plan or the organizational process improvement plan. Defining the evaluation criteria up front contributes to the objectivity of the evaluation activity as well as ensuring full coverage of data, information products, measurement activities, resources, and infrastructure. A good evaluation process will enhance an already adequate measurement program.

The evaluation tasks described in this section use predefined criteria. As shown in Figure 7-2, these criteria refine the PSM categories and measures to address the specific issues related to managing a measurement program.

Evaluation Task	***Evaluation Criteria (Measures)***	***PSM Category***	***PSM Issue***
Measures and Indicators	Accuracy Usability Reliability	Functional Correctness Usability Dependability - Reliability	Product Quality
Process Performance	Timeliness Efficiency Defect Containment Customer Satisfaction	Process Efficiency Process Efficiency Process Effectiveness Customer Feedback	Process Performance
Process Conformance	Audit Checklists	Process Compliance	
Process Capability	Reference Model	Process Compliance	

Figure 7-2. Evaluation Criteria Mapped to PSM Categories

The following sections describe the evaluation tasks and criteria in more detail.

2

Evaluate Measures and Indicators

The first task in the *Evaluate Measurement* activity evaluates the products of the measurement process: measures, indicators, and analysis results. The effectiveness of these measurement products should be evaluated with pre-defined criteria. The following criteria are adapted from ISO/IEC draft standard 15939, Software Measurement Process. Most of these criteria are subjective. Some criteria are specific to measures, indicators, or analysis results. These criteria are not necessarily independent of each other.

Measurement Products Use

This criterion measures the extent to which the project manager (or other measurement user) actually uses the analysis results and interpretations produced by the measurement process for. For example, how often does an inspection moderator use decision criteria provided by the measurement program to determine whether a reinspection should be performed? Most of the additional criteria described below have an influence on the use of measurement results.

Confidence in Measurement Results

This criterion determines if users of the information products have confidence in the measures, indicators, and interpretations incorporated in the measurement results. Greater confidence is achieved when the measurement users consider the measurement analysts to be competent and unbiased. Confidence also is increased if data providers are involved in the process, such as through regular feedback sessions.

Measurement Results' Fitness for Purpose

This criterion evaluates if measurement results effectively meet information needs. Not all indicators work well in all situations. A given measure may be easier or harder to collect under different circumstances, thus limiting the value of an information product. Confidence in a measurement product's fitness for purpose increases as evidence accumulates for its effectiveness in similar environments.

Fitness for purpose includes:

- The extent to which the measures and indicators actually measure what they are supposed to
- The extent to which predictive measures and indicators demonstrate the ability to forecast

To the extent that measurement results provide comprehensive and appropriate feedback relative to an information need, the results may be judged to be fit for their purpose.

Understandability of Measurement Results

This criterion gauges the ease with which the measurement user can understand indicators and interpretations produced by the measurement process. Measurement results are less likely to be used if they are difficult to

understand. Use of unfamiliar jargon may cause misunderstanding. Volume alone may be an obstacle to understandability - lengthy reports are less likely to be read carefully.

Satisfaction of the Assumptions of an Indicator Model

This criterion determines if an indicator's assumptions have been satisfied (e.g., data distributions, measurement scales, units of measure, sample size). Statistical techniques often rely on assumptions about the input data. Even simple numerical techniques usually depend on assumptions about what is being measured. Because violations of these assumptions usually occur in a specific context, an indicator model that depends on those assumptions should be avoided or, at least, interpreted with care. More sophisticated techniques generally make more complex assumptions and are more sensitive to violations.

Measurement Accuracy

This criterion examines whether the implementation of a measure conforms to its definition. Results may differ from what was intended because of problems such as systematic errors in the procedure, random errors that are inherent in the underlying measurement method, or poor execution of the procedure.

The actual human procedure or automated implementation of a measure may depart from the measure's definition. For example, a static analysis tool may implement a counting algorithm differently than the way it was originally described in the literature. Discrepancies may also arise from ambiguous definitions of items such as measurement methods, scales or units. Even good measurement procedures may be inconsistently applied, resulting in the loss of data or the introduction of erroneous data.

Subjective measurement methods depend on human interpretation. For example, the formulation of questionnaire items may mislead respondents about the question or may bias the responses. Clear, concise instructions help to increase the accuracy of surveys.

Increase the accuracy of measurement procedures by ensuring the following:

- The amount of missing data is within specified bounds
- The number of flagged inconsistencies in data entry is within specified guidelines
- The number of missed measurement opportunities is within specified thresholds (such as the number of inspections for which no data was collected)
- No inappropriate selection exists in the sampling process (such as surveying only satisfied users to evaluate user satisfaction, or evaluating only unsuccessful projects to determine overall productivity)
- All measures are well-defined, and those definitions are communicated to data providers

Poorly defined measures tend to yield inaccurate data. The reliability of the underlying measurement method (see below) may also limit the potential accuracy of a measurement procedure.

Measurement Reliability

This criterion measures the consistency of a measure's results when the measurement method is repeated. Random measurement error reduces reliability.

The reliability of a measurement method is evaluated on two factors:

- **Repeatability** - The degree to which the repeated use of the measure in the same organization, following the same measurement method, under the *same* conditions (such as tools or individuals performing the measurement) produces results that can be accepted as identical.

- **Reproducibility** - The degree to which the repeated use of the measure in the same organization, following the same measurement method, under *different* conditions (such as tools or individuals performing the measurement) produces results that can be accepted as identical.

Subjective measurement methods are generally less reliable than objective methods.

[This page intentionally left blank.]

3

Evaluate the Measurement Process

Defining good measures and indicators is not enough to ensure success. Even a measurement program based on appropriate, proven measures and indicators may be rendered ineffective by a clumsy and inflexible measurement process. Evaluate the measurement process from three perspectives:

- **Performance** - measurement of the inputs, outputs, and effects of the measurement process
- **Conformance** - comparison of the measurement process to a description of its intended use
- **Capability** - comparison of the measurement process to an external benchmark of process maturity

The following sections describe these perspectives in more detail.

All three dimensions should be assessed regularly. Quantitative performance evaluation requires a significant amount of process data. One year's worth of data is recommended. Audit conformance on a semi-annual basis; assess capability every two to three years at a minimum.

The measurement analyst usually evaluates measurement process performance, while independent groups typically conduct conformance audits and capability assessments.

3.1 Performance

The measurement process itself is subject to measurement. When evaluating the performance of the measurement process, consider the objective measures of output, schedule, and resource utilization as well as subjective feedback from the project manager and other users. The selected measures should address the information needs and issues of the people who perform and manage the measurement process. The data collected may also help to plan future measurement processes.

These four criteria are useful for evaluating the performance of a measurement process:

- Timeliness
- Efficiency
- Defect Containment
- Customer Satisfaction

The following sections describe the application of these criteria to the performance evaluation of a measurement process.

Timeliness

The measurement process should provide information to the project manager so that decisions can be made in time to affect the outcome of the measured activity. Depending on the phase of the project and the decision being supported, "timely" may mean daily, weekly, or monthly. The timeliness criterion for a measurement process corresponds to the PSM measurement category of Process Efficiency. A typical measure is the cycle time from collecting data to delivering analysis results. However, short cycle times are not as important as synchronizing delivery of information with the project manager's needs.

Measurement Process Efficiency

The benefits of the measurement process should outweigh the costs. The greater the benefit/cost ratio, the more efficient the process is. An alternative to the benefit/cost ratio is a return-on-investment figure, which includes inflation and the cost of money. Both approaches require information about benefits and costs. Determining the cost of a measurement process is generally less difficult than estimating the benefits (such as improved communication) that are often difficult to quantify.

Measurement process expenses include both investment and operational costs. Examples of investment costs include training, tool acquisition, and other setup costs. Operational costs primarily involve the labor to plan, collect data, analyze, and report results to project management. Increased automation is an investment cost that usually results in reduced operational costs for long-term measurement systems that can be reused from project to project.

No benefits can be realized from a measurement process unless action is taken based on the measurement results. The only way to separate the benefits of the action from the benefits of measurement is to clarify that the action could not have been taken without the insight from measurement. The benefits of a measurement process often result in cost avoidances, specifically through identifying and avoiding risks and problems so that their effects can be minimized. Benefits may be represented by the costs that would have been incurred if the problems had actually occurred. Some benefits may be intangible, such as improved employee morale resulting from the communication of project objectives and status.

Another method for determining benefits is to measure the cost of an operation at two points in time. First, measure the cost of the operation before the measurement process is implemented, and then measure it after measurement matures. Allocate credit for the change in operating costs between the measurement process and any other improvement initiatives that have been undertaken.

Defect Containment

The measurement process should minimize the introduction of erroneous data and results, while removing any defects as soon as possible. While Part 4 of the PSM Guide discusses some general strategies for verifying data and analyses, most organizations discover unique problematic areas. Gather feedback on defects discovered after analysis to develop additional strategies for minimizing defects in the measurement process.

The defect containment criteria map to the PSM category of Process Effectiveness. A typical measure would be the percent of data and analysis problems discovered prior to the presentation of the analysis results.

Customer Satisfaction

The project manager and other users should be satisfied with the information provided by the measurement process. Satisfaction results not only from the production of standard reports, but also from the response to

special information requests. Customer satisfaction indicates the user's assessment of the quality of the measurement information (see Chapter 2) relative to the performance of the measurement process in terms of timeliness, efficiency, and defect containment. Satisfaction may be affected by the user's expectations for quality and performance.

The customer satisfaction criterion maps to the PSM category of Customer Feedback. Customer satisfaction data is often obtained through a periodic survey of the project manager and other measurement users. The survey should ask measurement users to assess the value of the reports they receive based on their content, timeliness, and usability.

3.2 Measurement Process Compliance

A measurement process may be evaluated in terms of its compliance to a description of its intended function. This description usually is found in the measurement plan and supplementary documentation. The plan may be derived from the organization's documented process, such as standards and procedures, or adapted from an external reference such as the PSM Guide or ISO/IEC draft standard 15939. Compliance with the plan is usually determined by an independent audit. For example, audits may be conducted by the quality assurance group or as part of an ISO 9000 review of the organization's Quality Management System.

The following sections describe how to use the PSM Guide as the basis for auditing a measurement process by using the tasks within the PSM *Tailor Measures*, *Apply Measures*, and *Implement Process* activities. These sections present guidelines for evaluating these tasks. Each section includes criteria that could be included on an audit checklist. Tailor these checklist criteria to the intended local implementation of PSM before conducting an audit.

These criteria verify the process, adherence to the measurement procedures, and work products that were defined in the original measurement plan. The measures, data collection procedures, and common database management are examples of work products that may be defined in the measurement plan and reviewed regularly. Significant deviations from the planned process should be justified.

Auditing the *Tailor Measures* Activity

Figure 7-3 defines key audit considerations for each task within the *Tailor Measures* activity. Each item examines whether or not the principal outcomes of each task have been accomplished. The first task in the *Tailor Measures* activity is to *Identify and Prioritize Project Issues.* This means that a well-articulated set of project issues should have been produced and prioritized. During the next tailoring task, appropriate measures are selected for each of the high-priority issues. The final task produces a measurement plan that takes the project's technology and the process into consideration.

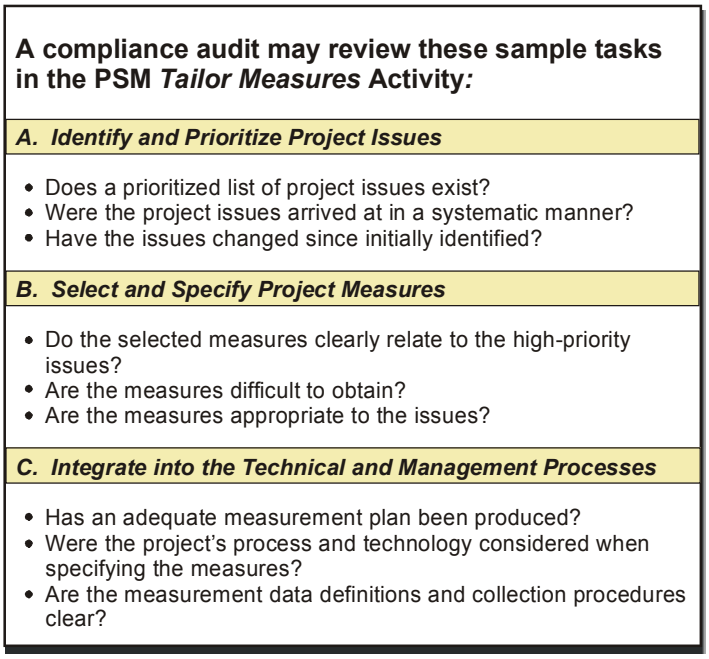

Figure 7-3. Sample Audit Criteria for the *Tailor Measures* Activity

The auditor should carefully examine the measurement plan. While this plan need not be a formal deliverable, certain key items should be covered (see Figure 2-10 in Part 2). Moreover, the information needs of a project usually change as the project progresses through its life cycle. The plan for a robust measurement process should change along with major milestones and project replans.

Auditing the *Apply Measures* Activity

Figure 7-4 defines key audit considerations for each task of the *Apply Measures* activity. These items explore whether or not the principal outcomes of each task have been accomplished. The first task in the *Apply Measures* activity is to *Collect and Process Data.* Inconsistent or inaccurate data may degrade the analysis results. Common causes of data problems include poor definitions, inconsistent formats, and data entry errors. The next task is to use the data to generate indicators for insight into the project's issues. The final task is to use those indicators and their interpretations to support decision making. A successful measurement process becomes a tool that the project manager relies on.

A compliance audit may review these sample tasks in the PSM *Apply Measures* Activity:

A. Collect and Process Data

- Is measurement data being generated as scheduled?
- Are the adequate resources provided to collect and manage data?
- Are the measures being consistently reported according to the procedures defined in the measurement plan?

B. Analyze Issues

- Are indicators being produced for all high-priority issues?
- Have all anomalies identified been investigated and the underlying causes described?
- Are the probable impacts of problems on project success estimated?

C. Make Recommendations

- Are indicators and interpretations provided to the project manager as scheduled?
- Are appropriate corrective actions identified?
- Are corrective actions tracked to closure and their impacts assessed?

Figure 7-4. Sample Audit Criteria for the *Apply Measures* Activity

The level of interaction between data providers and decision makers is a good indication of a measurement process' effectiveness. Data and analysis results should be reviewed with the data providers to ensure accuracy and consistency. Indicators and interpretations should be discussed with decision makers to answer any supplemental questions. Broad involvement in the measurement process across the organization ensures success.

Evaluating the *Implement Process* Activity

Successful execution of the *Tailor Measures* and *Apply Measures* activities depends on implementing an effective support structure for the measurement process. Figure 7-5 defines key audit considerations for each task of the *Implement Process* activity. Each item in the figure asks whether or not the principal outcomes of each task have been accomplished. The first task in the *Implement Process* activity is to obtain commitment. Commitment must be demonstrated in a tangible form. Once commitment is obtained, roles and responsibilities can be assigned. Next, appropriate resources are deployed. All involved parties must agree on the initiation schedule.

A compliance audit may review these sample tasks in the PSM *Implement Process* Activity:

A. Obtain Organizational Support

- Have the benefits of measurement been explained to all levels of the organization?
- Has senior management demonstrated support for measurement?
- Have privacy concerns been addressed explicitly?

B. Define Measurement Responsibilities

- Do all members of the organization understand their role in the measurement process?
- Are roles assigned to all participants appropriate to their backgrounds?

C. Provide Measurement Resources

- Have appropriate tools and staff been put in place to support measurement?
- Have all personnel been trained in their roles in the measurement process?
- Has a schedule for roll-out been developed and distributed?

Figure 7-5. Sample Audit Criteria for the *Implement Process* Activity

The ISO/IEC draft standard 15939 provides task descriptions that may be used instead of the PSM Guide to evaluate conformance of the measurement process. This standard only addresses the outcome (or the effect) of the process and its component tasks. An evaluation using this standard does not address the capability (maturity) of the measurement process, which is discussed in the next section.

3.3 Measurement Process Capability (Maturity)

ISO/IEC draft standard 15939 and this PSM Guide describe the basic tasks that must be performed in any measurement process. However, as a measurement process matures, additional tasks and more detailed task guidance may be added to improve process repeatability and performance. The organization may discover these enhancements through its own experience, or may adopt them as the result of a comparison to an external benchmark (or reference standard) of good practices. The Capability Maturity Model® Integrated Systems/Software Engineering and ISO/IEC 15504, "Information Technology Process Assessment," provide methods for evaluating the maturity (capability) of a measurement process.

This evaluation task uses one or more reference standards to assess the measurement process. The results of this evaluation may be used to identify improvements to the measurement process, to monitor improvement progress, or to select among alternative measurement processes. To illustrate this task, only the ISO/IEC 15504 method is discussed in this section. ISO/IEC 15504 provides a framework for assessing engineering and technical processes.

Conducting a good process capability assessment requires careful staffing and thorough planning. Regardless of how well defined the process assessment method is, the results of the assessment still involve subjective judgments by the assessment team. The team members must collectively possess the right blend of domain knowledge, engineering experience, and assessment skills for each specific evaluation.

A plan should be developed for each measurement process capability assessment. Adopting a well-defined method such as ISO/IEC 15504 facilitates planning. The contents of a typical assessment plan should include:

- Assessment purpose, defining the alignment with business goals
- Highest measurement process capability level that will be investigated
- Required inputs needed to perform the assessment
- Activities to be performed in conducting the assessment
- Resources and schedule for the assessment activities
- Information or data that will be collected during the assessment to determine the process capability or process improvement status
- Selection criteria and responsibilities of all participants in the assessment, including qualifications of the assessors
- Criteria for verification of assessment findings
- Description of the planned assessment outputs, including the ownership of the assessment outputs and any restrictions on their use

In most cases, the measurement process will be assessed as part of a larger assessment activity whose scope includes the life-cycle process. The following discussion focuses on the measurement aspects of such an assessment.

Most assessment methods involve reviewing documentation and interviewing practitioners. The measurement plan and related procedures provide information essential to the assessment. Other important artifacts include measurement reports and training materials. Practitioners are usually interviewed in functional groups, such as managers, measurement analysts, software engineers, and systems engineers.

The results of a capability assessment include strengths and weaknesses, as well as a summary capability level for the measurement process. Each assessment method provides a mechanism for determining a capability level based on observations from the documentation and interviews. The detailed assessment data is usually confidential so that individual sources cannot be identified. However, composite results should be provided to the assessed organization so that problems can be better understood and corrective actions can be taken.

[This page intentionally left blank.]

4

Update the Experience Base

The next task of the *Evaluate Process* activity is to update the measurement experience base with lessons learned from the evaluations and anecdotal information. The scope of lessons learned should encompass the measurement process as well as the measures and indicators. Consider both strengths and weaknesses. Capture things that did not work as well as things that did.

The experience base may range from a simple paper file to a separate organization that processes lessons for future use. Electronic bulletin boards are a common initial approach. As the measurement process matures, more sophisticated technology may be used to organize and disseminate the experience base.

A wide range of artifacts and observations may be in the experience base. Potential artifacts include:

- Measurement plans, policies, and procedures
- Definitions of measures and indicators
- Data verification techniques
- Customer satisfaction surveys
- Performance analysis reports
- Measurement process audit reports
- Measurement process capability assessment results

Types of observations include:

- Patterns symptomatic of specific problems
- Successful and unsuccessful corrective actions
- Implementation problems and results
- Tool evaluations

Lessons learned about measures and indicators may drive changes to the Issue-Category-Measure (I-C-M) table, especially if they were tailored to facilitate project or organizational use. Common results from the activity include new common issue areas, measurement categories, and measures.

Note that planned and actual project data accumulates in a project database that is usually physically separate from the experience base. However, archived copies of project data may be captured as components of the experience base as well.

[This page intentionally left blank.]

5

Identify and Implement Improvements

Some of the lessons gathered in the previous task can be implemented as improvements to the current project's measurement process. Others may be stored for the benefit of future projects. This task identifies specific improvements that can be implemented in the current project's measurement process. Consider the costs and benefits of potential improvements when selecting them. The actual implementation of selected actions may be part of the measurement plan's revision in the next iteration of tailoring. Or, the actions may be initiated as a parallel activity.

Changes that commonly affect the elements of a measurement plan include:

- Definitions of measures
- Analysis techniques
- Data collection and processing procedures
- Staffing and tool support
- Reporting and communicating

The effects of changes should be tracked to determine their effectiveness. Proposed measurement process changes that affect the project teams' activities, such as data generation, should be coordinated through the project manager.

[This page intentionally left blank.]