

# Measurement 101

Elizabeth C. L. Starrett

Software Technology Support Center (STSC)

*All too often, I read an article that uses terms or concepts I do not understand. The author, usually an expert on the subject, probably assumes the meaning of these terms is obvious and therefore does not explain them. If the terms and ideas appear frequently, I do not understand the article. In this article, I explain some common measurement information to clarify terms used in the measurement articles in this issue of CROSSTALK.*

This article is not for experts in measurement. It is simply a list of definitions, examples, and ideas that may be useful to the nonexpert. I will not dwell on how to set up a measurement program or suggest a measurement process because previous CROSSTALK articles have done well enough (some of them are listed at the end of this article).

## Definitions

When I first became involved with software measurement issues, I thought “metrics” and “measures” were synonymous, but because organizations and individuals define these terms different ways, they may or may not be synonymous. The following definitions used by the STSC are commonly used.

**Measure:** A standard or unit of measurement—the extent dimensions, capacity, etc.—of any thing, especially as determined by a standard; an act or process of measuring; a result of measurement [1]. Examples of measures include number of defects and source lines of code (SLOC).

**Metric:** A calculated or composite indicator based on two or more measures; a quantified measure of the degree to which a system, component, or process possesses a given attribute [1]. An example of a metric is defects per thousand SLOC (KSLOC).

Note that two or more measures make up a metric. Realize also that combining two or more metrics gives the information meaning. For example, while measuring 10 defects per KSLOC for a current project, how does one know if this is good, bad, or average? Comparing this metric with a previous project provides a baseline for the data and gives meaning to the metric.

## Cost/Schedule Control Systems

**Criteria (C/SCSC):** A Department of Defense (DoD) method established in 1967 to standardize contractor requirements for reporting costs and schedule performance on major contracts and to provide visibility of accomplishments on each contract. Other U.S. agencies have also adopted similar criteria [2]. Many DoD organizations refer to this method simply as “earned value.” Establishing a C/SCSC program tends to be complex and expensive, so it is important to know that earned value and C/SCSC do not have to be synonymous—an earned-value process can be implemented in an organization without implementing a formal C/SCSC-compliant function.

**Earned Value:** A measure of the value of work performed. Earned value uses original estimates and progress to date to show whether the actual costs incurred are within budget [3].

**Indicator:** A measure or combination of measures that provides insight into a software issue or concept [4]. For example, if an organization considers customer satisfaction to be an issue, defects per KSLOC found by the customer might be a good indicator of customer satisfaction.

**Normalize:** To cause to conform to a standard. Normalizing data is a process of dividing the numbers back into themselves to leave a percentage instead of actual numbers. For example, an organization that begins tracking its defects per KSLOC comes up with six defects per KSLOC. Every time the organization counts its defects per KSLOC in the future, that number is divided by 6. This gives a relative number that the organization can use to track whether it is improving. It also gives the organiza-

tion a number that releases minimal sensitive information if an outside organization sees it.

**Standard:** An accepted measure of comparison for quantitative or qualitative value. To continue the previous example, the standard number of defects per KSLOC delivered to the customer might be set at four defects per KSLOC. Future releases are compared to the standard of four, which provides a relative perspective to the developer and the customer. The use of an industry standard helps the organization understand how it fares in the marketplace. It is important to know that although formally designated standards are not available, informal industry standards have been established over the past several years based on the experience of recognized experts [5].

**Threshold:** A preset limit at which point action should be taken as a result of the data. Thresholds are established above or below the standard or both. They are often referred to as *upper* and *lower thresholds*. For example, the upper threshold of defects delivered to the customer may be six defects per KSLOC. If the contracting organization receives problem reports from the customer that show the defects in the delivered software are more than six per KSLOC, an investigation will begin to determine why there are so many errors, and the cause will be corrected.

The definition you choose for any of these terms is not as important as agreeing on common definitions for the entire organization. Ensure that everyone understands and uses those definitions consistently.

## Typical Measures

When starting a measurement program, the best approach is to reference the organization's strategic plan and use measurements that will indicate if the organization's goals are being achieved. Unfortunately, many organizations do not have a strategic plan. In this case, the first measurements should address the issues important to project managers and customers (most people involved have issues). The following five measures will often be part of the resulting measurement project.

### Size

The amount of software a project develops. The two most common ways to measure the size of software are SLOC and function points. One frequently asked question is, "How do I define SLOC?" There is no single answer for this; the most important thing is to clearly and consistently define it for the group being measured. Three common SLOC example definitions are

- The count the compiler gives when it compiles a program.
- Noncommented lines of code.
- "Any code that requires design, code, documentation, and test. This does not count debugging code that will not be delivered to the customer in the final product." [6]

Function points are sometimes used instead of SLOC and sometimes used in addition to SLOC. Function points measure software size by quantifying its functionality [7].

SLOC are preferred by many organizations because they can be easily counted, and everyone can understand a line of code. Function points are preferred by other organizations because the number will remain consistent across languages and platforms and because an accurate count can be determined at the requirements phase.

### Effort

Effort is the amount of work required to perform a task. Some example metrics for effort include

- Man-hours per phase of software development (requirements defini-

tion, design, code, and test are common lifecycle phases).

- Man-hours per defined set of requirements.
- Man-hours per project.

### Schedule

Schedule is the timing and sequence of tasks within a project [3]. The schedule may include tasks, milestones, activities, and phases required for the project.

It is important to note that the duration of a task is not necessarily the same as the effort involved. Two different projects may take the same amount of calendar days to complete but have differing amounts of effort. A project that takes one person working part time five days to complete has an effort of 20 man-hours, whereas another five-day project that employs two people full time will show 80 man-hours of effort.

### Cost

Some people consider cost as a combination of the effort and schedule, e.g., 20 man-hours per week for three weeks equals 60 man-hours. This method of calculation is flawed because it does not take into account the differing costs each organization has. A software customer may have two contractors with similar project requirements. The effort and schedule for the two contractors may be the same, but the cost will be different.

### Quality

People differ on what defines a quality product. Some may be happy with the quality if there are few defects. Others may care more about how user-friendly the software is; still others will be concerned about how easy the software is to maintain. Following are some examples of quality metrics.

- **Defects:** Defects per KSLOC, number of defects found per lifecycle phase, number of defects inserted per phase, cost to fix defects, impact of defects on delivered system, cause of defect insertion.
- **User friendliness:** Response time of system, capability of a system to recover from user errors.

- **Maintainability:** The ease or difficulty of keeping a system up to date and running. Different organizations have developed processes for scoring the maintainability of software products. The process developed by the Air Force Operational Test and Evaluation Center (AFOTEC) serves as one good example. This process includes a list of questions related to the software. A board is established that reviews the software and decides how well it meets the criteria of the questions. The board then scores the software on a scale of 1 to 6: "1" means that the software is nearly impossible to maintain and "6" that it is easy to maintain.
- **Rework:** Any effort in reaccomplishing work already deemed complete. Rework effort begins once a defect is found and continues until all the work required to obtain acceptance of the rework is complete [4]. "Already deemed complete" is the area of difference among organizations. Some consider this to be work deemed complete by the programmer, so any code change would be considered rework. Another organization does not consider changes to be rework unless they are changes made after software is released to the customer. The definitions follow the entire range within this time frame.

## Crime and Punishment

By law you must collect and act on measurements if you are a government organization. Rather than list every law here, following this article is a copy of Appendix A from the *Air Force Information Technology Investment Performance Measurement Guide*. The appendix is a list of the current laws that require measurements.

## Methods

There are numerous effective methods to implement a measurement program, but to list them all is beyond the scope of this article. However, I will mention the Practical Software Measurement (PSM) method. PSM is important because it is sponsored by the Joint Logistics Commanders (JLC) Joint Group on System

Engineering and is gaining acceptance among government organizations. (The JLC comprises members from each of the services who work on issues applicable to all parts of the DoD). PSM was developed as an aid to establish a measurement program. It currently includes a guidebook, training, and a software tool that implements the PSM process. *PSM: A Foundation for Objective Project Management* (typically called the PSM Guide) describes organizations at the beginning of the measurement process, indicating who should be involved and their responsibilities. The guide also suggests different issues that may affect a software project, what measurements would help track those issues, how to collect and analyze the measurements, and suggests how to act on the analysis conclusions. The guide is available from <http://www.psmc.com>.

### Existing Data

Organizations often ask me for example data of other organizations to which they can compare their data. Unfortunately, most organizations consider this data to be extremely sensitive and rarely release it to the public without normalizing it first. A limited amount of sample data can be accessed from the National Software Data and Information Repository (NSDIR). However, this repository has not been actively maintained for over a year, and the data is approximately two years old. The NSDIR can be accessed at <http://nsdir.cards.com/nsdir>.

### Closing

If you read an article that contains terms you do not understand or hear terms used that leave you confused, please feel free to contact us. We also welcome any editorial comments you may wish to send us on measurement or other software-related issues.

I also ask all measurement experts to keep in mind as you write an article,

give a presentation, or talk with customers that your audience is likely not expert—terms and ideas that are intuitive to you may not be intuitive to your audience. ♦

### About the Author

**Elizabeth C. L. Starrett** has been a software engineering consultant for the STSC for five years, where she helps clients improve their software processes. Her most recent duties include leading the STSC Measurement Team. She has spoken at the Software Technology Conference, the Data Reduction and Computer Group Conference, and has been published in *CROSSTALK*. Prior to joining the STSC, she worked for the Air Force with its supporting contractors to develop, document, and test data analysis and test support software for radar and the Peacekeeper missile. She has a bachelor's degree in electrical engineering from Utah State University.

Elizabeth C. L. Starrett  
OO-ALC/TISEC  
7278 Fourth Street  
Hill AFB, Utah 84056-5205  
Voice: 801-775-5555 ext. 3059  
DSN 777-9730  
Fax: 801-777-8069 DSN 777-8069  
Internet: starretb@software.hill.af.mil

### References

1. STSC Measurement Team, *Measurement Foundation Workshop*, 1994.
2. Fleming, Quentin W., *Cost/Schedule Control Systems Criteria the Management Guide to C/SCSC*, Probus Publishing Company, Chicago, Ill., 1992.
3. *User's Guide for Microsoft Project*, Microsoft Corporation, 1995.
4. *Practical Software Measurement: A Foundation for Objective Project Management*, Ver. 3.1, April 1998.
5. Jones, Capers, *Patterns of Software Systems Failure and Success*, International Thomson Computer Press, Boston, Mass., 1996.
6. Jensen, Randall W., "Estimating the Cost of Software Reuse," *CROSSTALK*, Software Technology Support Center, Hill Air Force Base, Utah, May 1997.

7. *Function Point Counting Practices Manual*, Dun & Bradstreet Software, Atlanta, Ga., 1994.

### Recommended Reading

#### Methods

1. Grady, Robert B. and Deborah L. Caswell, *Software Metrics: Establishing a Company-Wide Program*, Prentice-Hall, 1987.
2. STSC Measurement Team, *Software Metrics Capability Evaluation Guide*, Software Technology Support Center, Hill Air Force Base, Utah, October 1995.
3. Giles, Alan, "Measurement – The Road Less Traveled," *CROSSTALK*, Software Technology Support Center, Hill Air Force Base, Utah, April 1996.
4. Pitts, David, "Metrics: Problem Solved?" *CROSSTALK*, Software Technology Support Center, Hill Air Force Base, Utah, December 1997.

#### Sample Measures and Metrics

5. Webb, David R. and David Haakenson, "Making Metrics Work Miracles," *CROSSTALK*, Software Technology Support Center, Hill Air Force Base, Utah, August 1995.
6. Stark, George, "Maintenance Measures," *CROSSTALK*, Software Technology Support Center, Hill Air Force Base, Utah, July 1997.

#### Function Points

7. *Function Point Counting Practices Manual*, Dun & Bradstreet Software, Atlanta, Ga., 1994.
8. Heller, Roger, "An Introduction to Function Point Analysis," *CROSSTALK*, Software Technology Support Center, Hill Air Force Base, Utah, November 1995.
9. Garmus, David, "Function Point Counting in a Real-Time Environment," *CROSSTALK*, Software Technology Support Center, Hill Air Force Base, Utah, January 1996.

#### PSM

10. *Practical Software Measurement: A Foundation for Objective Project Management*, Ver. 3.1, April 1998.

*This reprint of Appendix A from the Air Force Information Technology Investment Performance Measurement Guide, August 1997 provides official directives with regard to metrics for Air Force and government organizations.*

## Appendix A – Guidance Documents

### Chief Financial Officers Act (CFOA) of 1990

The CFOA requires agencies to include performance measurement data in their annual financial statements.  
(<http://www.npr.gov/library/misc/cfo.html>)

### Government Performance and Results Act (GPRA) of 1993

The GPRA requires strategic planning and performance measurement in the executive branches of the government. Purposes are to improve federal management, congressional decision-making, service delivery, program effectiveness, public accountability, and public confidence in government. The GPRA requires agencies to develop strategic plans by September 30, 1997, for implementation in fiscal year 1999. The OMB (Office of Management and Budget) has mandated that the plans cover six years and be updated at least every three years. Stakeholders and customers will provide input into the strategic plans. Beginning in fiscal year 1999, agencies will develop yearly performance plans and set performance goals based on their strategic plans. Starting in March 2000, agencies will write annual performance reports, comparing actual performance to goals established in annual performance plans.  
(<http://www.hhs.gov/progorg/fin/gpraindx.html>)

### OMB Circular A-11, Part 2: Preparation and Submission of Strategic Plans

This circular provides executive guidance for preparing and submitting agency strategic and performance plans as required by GPRA.  
(<http://www.whitehouse.gov/WH/EOP/OMB/html/circulars/a011/toc97.html>)

### Federal Acquisition Streamlining Act (FASA) of 1994

The FASA contains specific requirements for federal agencies to “define the cost, performance, and schedule goals for major acquisition programs” and to monitor and report annually on the degree to which these goals are being met. Agencies must assess whether acquisition programs are achieving 90 percent of cost, performance, and schedule goals and, if not, determine whether to continue the program.  
(<http://thomas.loc.gov/cgi-bin/query/z?c103:S.1587.ENR:>)

### Paperwork Reduction Act of 1995 (PRA 95)

The PRA 95 intends to improve the quality and use of federal information; to minimize the cost to the federal government of the creation, collection, maintenance, use, dissemination, and disposition of information; and to ensure that information technology is acquired, used, and managed to improve performance of federal agency missions.

Per PRA 95, agencies must:

- Develop and maintain a strategic information resources management plan that shall describe how information resources management activities help accomplish agency missions
- Develop and maintain an ongoing process to:
  1. ensure that information resources management operations and decisions are integrated with organizational planning, budget, financial management, human resources management, and program decisions;
  2. in cooperation with the agency Chief Financial Officer (or comparable official), develop a full and accurate accounting of information technology expenditures, related expenses, and results; and
  3. establish goals for improving information resources management’s contribution to program productivity, efficiency, and effectiveness, methods for measuring progress toward those goals, and clear roles and responsibilities for achieving those goals.

(<http://www.os.dhhs.gov/progorg/oirm/pl104-13.txt>)

### OMB Circular A-130: Management of Federal Information Resources

This circular provides executive guidance on the management of federal IM/IT resources in compliance with PRA 95. Specific requirements include strategic IM/IT planning tying IT investments to agency mission accomplishment and cost/benefit analysis of IT systems throughout the system life-cycle.  
(<http://www.whitehouse.gov/WH/EOP/OMB/html/circulars/a130/a130.html>)

**Clinger-Cohen Act** (formerly known as Information Technology Management Reform Act [ITMRA]) of 1996  
The Clinger-Cohen Act directs that investments in IT support the mission, long-term goals and objectives, and annual performance plan of the department. It mandates that the Secretary of Defense implement performance measurement for all DoD IT programs, projects, and acquisitions.  
(<http://www.dtic.mil/dodim/cohen.html>)

**OMB Circular A-11, Part 3: Planning, Budgeting, and Acquisition of Fixed Assets**

This circular provides executive guidance on planning, budgeting, and acquisition of fixed assets, specifically IT and NSS-IT, in accordance with GPRA and Clinger-Cohen Act. It requires agencies to identify baseline goals for cost, schedule, and performance for all proposed and ongoing acquisitions, and provides guidance on reporting compliance with these goals to OMB.  
(<http://www.whitehouse.gov/WH/EOP/OMB/html/circulars/a011.toc97.html>)

**Executive Order 13011, Federal Information Technology**

This order implements the provisions of Clinger-Cohen Act in the executive branch. Besides the specific provisions of Clinger-Cohen Act, the order establishes the Federal CIO Council; creates the Government Information Technology Services Board and the Information Technology Resources Board; and provides additional guidance on the roles of agency CIOs and the use of performance measurement in evaluating IT investments.  
(<http://www.npr.gov/library/direct/orders/27aa.html>)

**Executive Office of the President, Evaluating Information Technology Investments – A Practical Guide (OMB Information Technology Investment Guide), November 1995**

Provides an analytical framework for linking IT investment decisions to strategic objectives and business plans in the federal organizations.  
(<http://www.whitehouse.gov/WH/EOP/OMB/infotech/infotech.html>)

**GAO, Executive Guide: Effectively Implementing the Government Performance and Results Act, (GAO/ GGD-96-118), June 1996**

Identifies key steps needed to implement GPRA, along with key steps that agencies need to take toward its implementation.  
(<http://www.gao.gov/special.pubs/gpra.htm>)

**DoD, Guide for Managing Information Technology (IT) as an Investment and Measuring Performance, February 1997**

The guide summarizes the DoD position on IT performance measurement and presents a framework for managing information technology programs as investments rather than as acquisitions.  
(<http://www.dtic.mil/c3i/cio>)

**DoD, Information Technology Management (ITM): Supporting National Defense (ITM Strategic Plan), Version 1.0, March 1997**

(<http://www.dtic.mil/c3i/cio>)

**Air Force Information Resources Management VISTAS (Air Force Information Resources Management Strategic Plan)**

(<http://www.cio.hq.af.mil/docs/vistas.htm>)

**DoD 5000.2-R, Mandatory Procedures for Major Defense Acquisition Programs and Major Automated Information System Acquisition Programs**

AFI 10-601, Mission Needs and Operational Requirements Guidance and Procedures.  
(<http://web7.whs.osd.mil/dodiss/publications/pub2.htm>)

This document was prepared for Arthur Money, chief information officer (CIO), U.S. Air Force by Andrulic Corporation.  
For information concerning the project, contact

James Brown  
CIO Support Directorate

AFCIC/ITIM  
1250 Air Force Pentagon, Room 4A1088E  
Washington, DC 20330-1260  
Voice: 703-697-3492  
Fax: 703-614-4471, -6346  
E-mail: brownjd@af.pentagon.mil