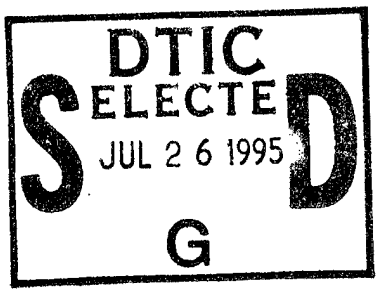
Technical Report
CMU/SEI-94-TR-016
ESC-TR-94-016

Carnegie Mellon University
Software Engineering Institute

A DoD Software Measurement Pilot: Applying the SEI Core Measures

James A. Rozum
William A. Florac

May 1995



Technical Report
CMU/SEI-94-TR-016
ESC-TR-94-016
May 1995

A DoD Software Measurement Pilot: Applying the SEI Core Measures

James A. Rozum

William A. Florac

Software Process Program

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

This report was prepared for the

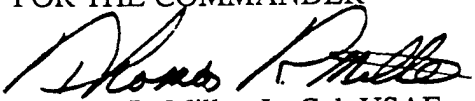
SEI Joint Program Office
HQ ESC/ENS
5 Eglin Street
Hanscom AFB, MA 01731-2116

The ideas and findings in this report should not be construed as an official DoD position. It is published in the interest of scientific and technical information exchange.

Review and Approval

This report has been reviewed and is approved for publication.

FOR THE COMMANDER



Thomas R. Miller, Lt Col, USAF
SEI Joint Program Office

This work is sponsored by the U.S. Department of Defense.

This document is available through Research Access, Inc., 800 Vinial Street, Pittsburgh, PA 15212.
Phone: 1-800-685-6510. FAX: (412) 321-2994.

Copies of this document are available through the National Technical Information Service (NTIS). For information on ordering, please contact NTIS directly: National Technical Information Service, U.S. Department of Commerce, Springfield, VA 22161. Phone: (703) 487-4600.

This document is also available through the Defense Technical Information Center (DTIC). DTIC provides access to and transfer of scientific and technical information for DoD personnel, DoD contractors and potential contractors, and other U.S. Government agency personnel and their contractors. To obtain a copy, please contact DTIC directly: Defense Technical Information Center, Attn: FDRA, Cameron Station, Alexandria, VA 22304-6145. Phone: (703) 274-7633.

Table of Contents

List of Figures, Checklists, and Reporting Forms

Figures

Checklists Used (Appendix B)

Reporting Forms (Appendix C)

Acknowledgments

Of course, an effort of this magnitude would not be possible without many contributors. Most notable of those contributions is from Mr. Ken Kelley of DISA/CFSW[1] who had the vision to initiate and sponsor the pilot and who led us through it. From the SEI, Ms. Anita Carleton and Mr. Archie Andrews made significant contributions to the preparation of this report and the pilot effort in general, and we sincerely thank them.

We gratefully acknowledge the support of the pilot's site champions who participated in this effort and, without whom, this report and the pilot would not have been possible:

Sheila Dennis; Defense Finance and Accounting Service; Denver, CO

Sandee Guidry; Defense Finance and Accounting Service; Pensacola, FL

Art Fluellen; Defense Logistics Supply Center-E; Battle Creek, MI

Harry Kelso; Navy Fleet Material Support Center; Mechanicsburg, PA

Sue Gipson; OL-AD MSC/SO; Hill AFB, UT

Eileen McGraw; Defense Supply and Acquisition Center; Columbus, OH

Louis Damiano; Defense Supply and Acquisition Center-V; Philadelphia, PA

John Culbertson; HQ CSC/SDQP; Tinker AFB, OK

Pamela Newenstyn; Navy Management Systems Supply Office; Chesapeake, VA

We are also grateful to the site champion's management who supported them in this effort, and, where applicable, the champion's predecessor.

Thanks go to those who supplied data for this technical report or reviewed various versions, including Bob Rathbun and Mike Falat from DISA/CIM; Tim Oncale from SAIC; and Wolfhart Goethert, James Herbsleb, and Andrew Huber (a resident affiliate from Data General at the time) from the SEI. Special thanks go to Laura Malone and Suzanne Couturiaux from the SEI who helped edit this technical report.

---

[11] The pilot was sponsored and managed by the Defense Information Systems Agency Center for Information Management (DISA/CIM) originally. At the end of the pilot initiative, DISA went through a reorganization and the portion of CIM that sponsored the pilot moved into a new division, Center for Software (CFSW). Throughout this report, therefore, the sponsoring and managing agency at DISA is referred to as DISA/CFSW.

Executive Summary

In January 1992, the Software Engineering Institute (SEI) was requested to make recommendations for a basic set of software measurements to support the software measurement initiative under the software action plan (SWAP) sponsored by the Department of Defense (DoD) software technology strategy. The recommendations, published in September 1992, included materials and guidelines for a set of basic measures (size, effort, schedule, and quality) that were intended to help the DoD plan, monitor, and manage its internal and contracted software projects [Carleton]. In the fall of 1992, a policy memorandum addressing software measurement for information management (IM) was proposed as a key element for improvement of software development and maintenance within the DoD. Although the policy memorandum had not been issued, the Defense Information Systems Agency, Center for Software (DISA/CFSW) was tasked to conduct a pilot study on the core measures.

The DISA/CFSW sponsored and managed the software measurement pilot. The purpose of the pilot program was to gain an understanding of the key issues that need to be resolved before a DoD-wide metrics program is initiated. Specifically, the intent was to use the lessons learned from the pilot to help develop a set of software measurement program guidelines for the DoD. The issues of concern to the DoD are reflected in the following objectives:

- Assess the ability of organizations to collect and use the core measures.
- Determine the applicability of common definitions of the SEI core measures across application domains and multiple sites.
- Gain an understanding of the cost of implementing and sustaining a viable metrics program.
- Determine the appropriate analysis and reporting of the measurement information at an organization and a corporate level.
- Evaluate the effectiveness of the SEI core measurement checklists for development and maintenance and recommend improvements.
- Develop metrics program guidelines for DoD implementation.

From the pilot program, valuable lessons were learned in the areas of policy guidance, training, data collection guides, limitations on common definitions, the impact of new or changed processes to project or sites, management participation, and the usefulness and role of the SEI definition checklists. These were topics that the DISA/CFSW pilot set out to understand in order to determine the ability of DoD organizations to implement software measurement-driven policies; from that standpoint, the pilot program was successful.

The pilot program also obtained information that provided some insight into the relative cost of implementing a measurement program. While the figures gathered probably do not fairly represent the absolute costs, they may be representative, in a relative sense, of the relationship between the cost of infrastructure improvements and the number of sites and projects.

Pilot Implementation

DISA/CFSW identified a number of representative sites that volunteered to participate in the pilot effort. DISA/CFSW trained and prepared the site champions to collect the SEI core measures, provided sites with consulting support, created a data repository to allow reporting and analysis of data, and sponsored the report on the findings of the pilot study.

A total of 31 projects were selected by the 9 sites. Typically, there were 3 to 4 projects per site, but one site had 6 projects and another had only one. The projects were primarily information management systems, but included a wide range of application areas. The projects were also mostly in the maintenance phase of the software life cycle. The 31 projects did not include any embedded systems projects or projects under acquisition contract with commercial firms.

Each site appointed a site champion as the point-of-contact for software measurement who acted as the liaison between the pilot and the projects. Periodic meetings with the pilot site champions served as the primary pilot management mechanism. At the meetings, the site champions reviewed the status of their respective site's effort to implement the required measures. The meetings also provided a method for pilot participants to

- Raise issues confronting them and identify areas where they required assistance.
- Share problems and solutions for implementing the measures.
- Determine status and assess progress of each of the sites in establishing measurement processes, and collecting and using the data.

Observations and Lessons Learned

Many observations, issues, and problems encountered during the pilot effort were typical of any project or organization initiating a software measurement program. This was not entirely unexpected, and numerous reports have been published citing the lessons learned from such an effort. This report, however, limits its observations and lessons learned to those consistent with the objectives of the pilot program. Many of these issues and observations are likely to exist at other DoD sites, and the lessons learned may affect how the DoD implements a standard and/or policy requiring software measurement or reporting of software data.

Two overriding lessons emerge from the totality of the lessons learned during the pilot study. They specifically relate to the establishment of a measurement process across multiple sites, domains, and application types:

- Organizational, communication, and personnel issues transcend the entire process and are just as significant, if not more so, than the technical issues related to software measurement. Paying attention to these issues will not guarantee success, but ignoring them will certainly result in a failed or less than satisfactory measurement process. This may be the most important lesson of all.
- Those responsible for the initiation and implementation of a measurement process across multiple sites must be aware that every action, and inaction, has a ripple effect that is both broad and deep across the organizational units. For example, when data are collected in response to a policy or management request, every action regarding how the data are used will affect the measurement process. Likewise, taking no action in response to the collected data will also have profound effects on the measurement process.

Other observations addressed in detail in the report include the following:

- *Collection and use of common measures:* There were many observations regarding the pilot objective to determine the impacts on an organization that is required to collect software measurement data that might be specifically defined in a policy. Observations dealt with how an organization might have to change one or more of its software processes to collect the data and the level of detail that would be needed to describe how each organization must collect and report the data.
- *Cost of a measurement program:* Each site was required to collect the amount of time the site champion and other site personnel spent in any aspect of trying to implement the required measures. The cost varied largely by site and appeared to be influenced strongly by the site's current infrastructure and software process. The costs of implementing the measurement program rose moderately as additional projects collected the measures. As a site's measurement process developed and matured, the overall costs seemed to decrease.
- *Analysis and reporting of measurement data:* There were several observations that were related to this area. A critical issue that arose, and that seems to have an impact on the ability of a site to sustain a measurement program, was the attention the site's management gave to the data collected. Those sites where managers reviewed the data and began to use the data in their own business processes appeared to be successful in implementing the measurement program. An additional issue of importance was how the measurement data were used. Those sites that could show how the data would be used prior to collection, and then demonstrate the use of data with regard to the stated purposes, appeared to overcome pockets of resistance to collecting the data.

- *Organizational issues:* It appears that to implement a common measurement program across an entire site or across multiple sites, the site champion must have the responsibility and authority to implement the data collection at the site. Also, linking the data collected to organizational goals communicates why certain data items are being collected and, thus, was an important aspect to the pilot.

Conclusions and Recommendations

The conclusions and recommendations that follow are repeated from the body of the technical report. Although we believe there are bits of value in the report to many different audiences, we feel the conclusions and recommendations are the key pieces of information needed by executives who may read this summary.

A large organization with multiple sites can write and implement a policy on software measurement that uses the SEI core measures as the basic units of measurement. Given what we have observed during this pilot, however, such a policy should be carefully crafted around specific purposes or objectives that the organization hopes to achieve with the measurement data collected.

For example, a policy may be implemented similar to the Hewlett Packard (HP) 10X program. In that program, HP issued a policy that its organizations would improve quality by a factor of ten. The effectiveness of any such policy issued will depend on the continual reinforcement of the policy objectives and regular review of the measures used to measure progress and performance toward meeting the policy objectives.

There are several methods for implementing such a policy. Two of those methods include the following:

- Issue a policy or directive at the headquarters level (the Office of Secretary of Defense within the DoD) that establishes one or more general objectives and requires the various subordinate organizations to measure progress towards the objectives in a specific manner.
- Issue a policy or directive at the headquarters level that requires subordinate organizations with significant involvement in software acquisition, development, or maintenance to establish objectives relative to their own organizational goals and to measure and report progress toward those goals. In this alternative, the affected subordinate organizations become the sponsoring organizations for specific policies that identify goals and objectives directly related to the organization and identify the means by which progress will be determined.

As evidenced by this pilot program and other efforts documented in the literature, to implement such policies successfully, it is essential to provide sufficient provisions for funding, staffing, training, and establishing a measurement process. It is recommended that

the organization issuing the policy provide guidelines and standards that address these issues relative to the policy implementation, and identify supporting resources to provide guidance and assistance to the subordinate organizations.

However the policy is actually issued, when a policy includes the requirement to collect and report software measurement, the organization issuing the policy needs to address the following:

- Define measurements to be adhered to by those reporting data. The definition can be articulated using a combination of SEI checklists describing what is to be included and excluded in the data reported and textual descriptions to further describe aspects of data items that cannot be described by a checklist.
- Make use of the SEI checklists to communicate the results. The policy should use the checklists to define the data to be reported. Sites collecting the required data will need to describe how they meet the policy's definition and possibly tailor or elaborate the policy definition with respect to the organization's software process.
- Develop, create, and provide a measurement process guide or handbook. The process guide should describe how and what can be tailored by a site to meet the requirements in the guide. The contents of a process guide should include
  - how the sponsoring organization will use the data to determine performance toward the objectives described in the policy;
  - the definition of the data contents to be collected (i.e., using the SEI checklists, describe what is to be included and excluded);
  - how the data are to be reported by a site to the organization, along with accompanying reporting forms for accomplishing this;
  - how a site can access the data stored at the organization level; and
  - how a site can expect to get feedback from the organization regarding its data.
- Identify the training and assistance available to the sites and how a site's point of contact can contact the organization issuing the policy. Recognize and provide the training that is needed for those developing the measurement process and also for those that will be expected to use the measurement results in their own software processes.
- Keep the reported data in a repository operated by the organization issuing the policy. Access may be unlimited, but the confidentiality of the data must be strictly enforced. That is, the data itself should be completely anonymous to those outside of the organizational chain. The database should include certain validity checks on the data prior to accepting the data into the repository.
- Require data to be reported electronically. The electronic transmission could be completed using a supplied template in the form of magnetic media or via telecommunication vehicles. This would require that the organization issuing the policy develop a reporting template and harmonize that template with the definitions and the database.

- Conduct pilot programs before attempting broad implementation.
- Recognize that not all sites will be able to comply with the policy immediately and a start-up period of time will be needed to implement the measurement process across the organization. The amount of time needed to comply will be a factor of the number of sites and projects involved. In a large, distributed organization, a year to bring all sites or projects into compliance with the new policy might not be unreasonable.

A DoD Software Measurement Pilot: Applying the SEI Core Measures

Abstract. A pilot effort was initiated by the U.S. Department of Defense (DoD) and led by the Defense Information Systems Agency (DISA) to assess the issues and effort involved in implementing a software measurement program across multiple sites and projects. The pilot was conducted involving multiple DoD organizations and projects from varying application domains. The objectives were to assess the ability of an organization to collect and use the core measures, determine the appropriate analysis and reporting of the measurement information at an organization and a corporate level, determine the applicability of common definitions of the SEI core software measures, evaluate the effectiveness of the SEI core measure checklists for development and maintenance, and develop metrics program guidelines for DoD implementation. This technical report discusses the observations and lessons learned from the pilot effort and makes recommendations regarding software measurement implementation across a large organization.

1. Introduction

This technical report documents the observations and lessons learned during a pilot program conducted to assess the issues and relative effort involved in implementing an organizational software measurement program across multiple sites and projects. The Defense Information Systems Agency, Center for Software (DISA/CFSW)[2] sponsored and managed the software measurement pilot. Earlier, the SEI recommended a set of core measurements to the Department of Defense (DoD) as a basic set of software measures to be used to help plan and manage the acquisition, development, and support of software systems. DISA/CFSW sought the support of SEI technical personnel experienced in the development of the core measurements, frameworks, and training modules to provide technical consultation to the pilot measurement effort.

---

[2] The pilot was sponsored and managed by the DISA Center for Information Management (DISA/CIM) originally. At the end of the pilot initiative, DISA went through a reorganization and the portion of CIM that sponsored the pilot moved into a new division, Center for Software (CFSW). Throughout this report, therefore, the sponsoring and managing agency at DISA is referred to as DISA/CFSW.

1.1 Audience and Report Organization

The primary audience for this technical report is those involved with implementing a software measurement program across multiple sites. We strongly believe that the observations and lessons learned are also very appropriate for those implementing a measurement program at a single site, but where data are expected to be collected across many projects, all of which may be using processes that vary in some way.

An executive summary is provided for those who need to implement a policy that directly requires the collection of specific software measurement data.

Readers interested only in browsing the specific lessons learned may go directly to Chapter 4 for an enumeration and discussion of the key lessons learned. Chapter 3 discusses our observations from the pilot and key issues and problems related to those observations. Those readers wanting more background and an overall description of the pilot effort, as well as an overall profile of the pilot projects, should read Chapters 1 and 2. Chapter 2, which provides an overview of the projects involved in the pilot, will also help readers of Chapter 3 put some of the observations into perspective. Chapter 5 contains the recommendations and conclusions. Appendices contain copies of the data definition checklists, document comment forms, and pilot data report forms that were used by the pilot.

1.2 Background

In January 1992, the Software Engineering Institute (SEI) was requested to make recommendations for a basic set of software measurements to support the software measurement initiative under the software action plan (SWAP) sponsored by the Department of Defense (DoD) software technology strategy. The recommendations, published in September 1992, included materials and guidelines for a set of basic measures that were intended to help the DoD plan, monitor, and manage its internal and contracted software projects [Carleton]. In particular, the SEI established frameworks for specifying or describing definitions for four basic measurement topics: size, effort, schedule, and quality.[3]

In the fall of 1992 a policy memorandum addressing software measurement for information management (IM) was proposed as a key element for improvement of software development and maintenance within the DoD. The draft policy memorandum stated:

---

[3] The frameworks for defining the measures were published as SEI technical reports. See [Park] for defining software size as measured by lines of code. For defining measures of effort (staff-hours) and schedule (milestone and deliverable dates and work status, see [Goethert]. For defining measures that use defect and problem report data, see [Florac].

All IM software-intensive programs will collect appropriate software metrics.

Initially the DoD metric set will consist of the Software Engineering Institute's developed 'core' metrics (size, effort, schedule, and quality).

To validate metric concepts, DISA will establish a pilot metrics collection program to begin collecting software metrics....and to evaluate the impact of metrics collection activities on software development processes.

Although the policy memorandum had not been issued, the DISA/CFSW was tasked to conduct the pilot study, with the intent that the policy memorandum would be eventually issued and that information from the pilot program would be used to construct that policy. In early 1993, the (now) DISA/CFSW Software Systems Engineering Directorate, was given the task to conduct the pilot.

The division chief developed a plan and contacted DoD central design activity (CDA) sites as potential metric pilot sites. The division chief also worked with the SEI to obtain help in implementing the pilot study and to provide consulting assistance to the pilot sites. Specifically, the SEI was tasked to

- Provide training and assistance to the pilot sites.
- Support DISA/CFSW with coordination and planning of pilot working group meetings.
- Participate in the pilot working group meetings by providing technical guidance and recommendations.
- Support DISA/CFSW in the analysis, collection, and validation of pilot project data.
- Prepare a post pilot review and analysis.
- Make recommendations to DoD for policy formulation.

1.3 Objectives

1.3.1 Objectives for the Pilot Program

The purpose of the pilot program was to gain an understanding of the key issues that need to be resolved before a DoD-wide metrics program is initiated and to use the lessons learned from the pilot to help develop metrics program guidelines for DoD implementation. The objectives that follow were crafted to gain an understanding of the issues:

- Assess the ability of an organization to collect and use the SEI core measures.
- Determine the applicability of common definitions of the SEI core measures across application domains and multiple sites.

- Gain an understanding of the cost of implementing and sustaining a viable metrics program.
- Determine the appropriate analysis and reporting of the measurement information at an organization and a corporate level.
- Evaluate the effectiveness of the SEI core measure checklists for development and maintenance and recommend improvements.
- Accrue information that can be used to develop metrics program guidelines for DoD implementation.

1.3.2 Objectives Related to the Measures Collected

In keeping with well-founded and widely known principles of establishing measurement processes [Basili], the pilot program established goals or issues that could be used to determine the measurement data requirements, i.e., define the measurements. The goals and issues, though, were those commensurate with the pilot, i.e., not goals relative to project management. Early in the planning for the pilot, consideration was given to the notion of using DoD goals or issues (e.g., improved productivity, quality, cost and schedule commitments) to establish software measurement requirements, thereby giving reason and purpose to the collection of data across multiple sites.

For a number of reasons (discussed in Chapter 3), some months after the start of the pilot, this was replaced by the notion of collecting data that would support the site projects in their project management and process improvement activities. Since many of the sites had software process improvement initiatives underway, this approach had the potential advantage of being helpful with these initiatives; and, just as importantly from DISA's perspective, allowed the pilot to proceed without time-consuming sessions to set goals and define issues within the DoD management environment. The requisite that the sites actually use the data was not an explicit pilot objective and not stipulated to be a requirement of pilot participants; rather, it was viewed as an opportunity for the sites to take advantage of the data as they saw fit.

DISA's role in this pilot effort was primarily that of a facilitator and service provider by organizing meetings between the site champions, engaging the SEI to provide technical consulting assistance, and providing a common data repository for the collected data with the support of Science Applications International Corporation (SAIC), a software engineering technical assistance (SETA) contractor to DISA/CFSW.

1.4 Description of the Pilot Program

DISA/CFSW identified a number of representative CDA sites that volunteered to participate in the pilot effort. DISA/CFSW trained and prepared the site champions to collect the SEI core measures, provided sites with consulting support, created a data repository to allow reporting and analysis of data, and sponsored this report on the pilot study. Subsequent sections in this chapter describe the pilot effort, the pilots sites and projects, and the activities and responsibilities of the various pilot participants. This description provides context and perspective that may be helpful in appreciating the observations, conclusions, and lessons learned in subsequent chapters.

1.4.1 Site Selection

Initially the pilot plan was to include at least five sites representing Department of Defense organizations and the three services. Seven volunteer sites were selected by the time the site preparation meetings were held, and ultimately grew to nine physical sites representing four organizational groups. Each site had at least one and as many as six participating projects. The organizations and sites participating included

- Defense Finance and Accounting Service (DFAS) sites (Denver and Pensacola).
- U.S. Air Force bases (Hill AFB and Tinker AFB).
- U.S. Navy sites (Navy Management Systems Support Office, Chesapeake, Va. and Fleet Material Support Office, Mechanicsburg, Pa.).
- Defense Logistics Agency (DLA) sites (Columbus, Oh.; Philadelphia, Pa.; and Battle Creek, Mich.).

1.4.2 Site Resources and Responsibility Expectations

DISA/CFSW would require an individual to be assigned by the site to be the site's metrics champion for the pilot program. DISA/CFSW provided a profile to the sites of the type of employee that was desired to fulfill the champion role. According to this DISA/CFSW profile, the site champion would have the following characteristics:

- Project manager experience.
- Senior technical person.
- Respected by peers and trusted by project managers.

The following were the expectations or responsibilities of a site champion:

- Work full time on the pilot program for the first two to three months establishing data collection processes, collecting the data from the participating projects, and forwarding the data to the data repository provided by DISA/CFSW.

- Assist the projects by working with them to collect, analyze, and use the measurement data from the project.
- Attend the planned training session(s) and attend periodic working group meetings (held at one of the pilot sites) to review progress and share information with the other pilot sites. The periodic working group meetings were to be held once per month for the first three to four months and every two to three months afterwards, until the end of the pilot program.
- Record the time that they and other site personnel expended on all measurement aspects related to the pilot program.

1.4.3 Preparation for Pilot Sites

DISA/CFSW conducted a four-day workshop for the pilot site champions at the start of the pilot effort. The objective of the workshop was to define, in detail, the data to be collected and reported by the sites for each of the SEI core metrics, i.e., size, effort, schedule, and quality.

A full-day session was devoted to providing the site champions with the basics of software measurement, including a review of the goal-question-measurement (GQM) paradigm [Basili], an introductory session on the SEI core measures, and a discussion of typical software metrics describing what they are, how they are used, how they are obtained, etc. On the following three days, the site champions attended a detailed tutorial on the SEI core measurement checklists, which included work sessions using the checklists to establish common definitions for each of the core measurements. During the training, the pilot site champions also defined common data collection and reporting forms.

Several guidelines for the pilot effort were suggested by DISA/CFSW and agreed upon by the site champions:

- Reach agreement on definitions that are common to all sites.
- Define data that are currently available or collectable.
- Use the SEI checklists to define the data.

After conducting a detailed tutorial for each of the respective checklists, DISA/CFSW offered a completed checklist to be used as a starting point for the definition of each of the measures (size, effort, schedule and quality). The site champions collectively reviewed the DISA/CFSW proposed checklist definitions and made modifications to the proposal that were believed to be both common and consistent with the data available at each site.

After defining the measurements, the site champions designed data reporting forms (using the forms described in the SEI checklist reports as a guide) corresponding to the measurement definitions. Both DISA/CFSW and the site champions realized that the pilot's

common measurement definitions and corresponding report forms needed to be tested in practice and were subject to modification. Specifically, the data availability had to be reviewed at each site, and the defined measurements had to be useful for the project manager.

Some months after the initial training, SEI advisors, after analyzing the data being collected, made several recommendations for modifying the definitions to better address common project management and process improvement objectives. The completed SEI checklists used by the pilot sites throughout the remainder of the pilot are found in Appendix B.

The initial reporting forms were eventually modified to clarify the reported data and to better meet the needs of the sites for supporting project management. Copies of the final reporting forms that were used by the pilot sites are included in Appendix C.

In addition to defining the core metrics for their pilot, the site champions also agreed to report the number of staff-hours that they and the project teams expended on measurement pilot-related efforts. The sites also agreed to supply one-time project profile data for each project.

1.4.4 Data Collection

The pilot site champion was responsible for collecting the software measurements and reporting them to DISA/CFSW for entry in the pilot data repository. The data were collected monthly and reported using manually prepared data report forms. It was up to the respective site champions to devise a data collection process for each of the participating projects, based on the software process and project management process at each site. Typically, the site champion met with each of the project managers involved in the pilot study and, using the checklists and data report forms created at the workshop, established a process to provide the necessary data. Frequently the site champion had to meet with personnel outside of the project (e.g., labor accounting and configuration management organizations) to make arrangements to obtain the required effort and size data. As the pilot progressed, the site champions were advised to document the process they had devised, and the DISA/CFSW consultants (SEI) provided an outline of the data collection process for guidance.

1.4.5 The Data Repository

DISA/CFSW had a data repository created to store, aggregate, and enable the analysis of the collected data. The data repository was implemented using a commercial-off-the-shelf (COTS) database management system on a personal computer running Windows 3.1 and MS-DOS. The DISA/CFSW SETA contractor was tasked to design and implement the

database, the data entry screens, the data analysis graphing tools; and perform the data entry duties. As the data collection process matured and the data repository became functionally complete, a copy of the database with only that specific site's data was made available to each site.

The database consisted of 6 major tables: project profiles, project effort, size, schedule, defects, and site pilot measurement effort. Each of the tables contained the data elements requested on the data collection forms for each project and spanned 14 months of the pilot program.

1.4.6 Reporting and Analysis of Project Data

A status briefing was given jointly to all attending CDA managers at a meeting hosted by the SEI approximately half-way through the pilot program. Also, DISA/CFSW provided periodic reports of the collected information for use by each site, preserving confidentiality of the site and projects involved.

As the site data became sufficient for trend analysis, DISA/CFSW performed analyses on available site data, and forwarded the analyses to the site champions to illustrate the ways in which the measurements might be of help to the managers. Also, DISA/CFSW engaged consultants who used a commercially available database of information to compare productivity and staffing characteristics of certain pilot projects (those who agreed to the analysis and had submitted sufficiently complete data) with similar projects in the database having the same project size, effort, and domain characteristics.

2. Overview of the Pilot Site Projects

At the time this report was written, a total of 31 projects from the 9 sites had participated in the DISA/CFSW measurement pilot. Typically, there were 3 to 4 projects per site, but one site had 6 projects and another had one. The projects were mostly information systems applications, but included a wide range of application areas. The following list of applications supported by the projects indicates the variety among the pilot project applications:

- Procurement
- Contract management
- Communications
- Weather
- Aircraft maintenance
- Air traffic control
- Stock and inventory management
- Personnel management

As shown in Figure 2-1, the list does not include embedded systems projects or projects under acquisition contract with commercial firms. Therefore, the pilot program did not experience or reveal any issues that might be associated independently with these types of programming efforts.

2.1 Project Activity

The pilot projects were primarily engaged in maintenance activities. Since the release dates for the major enhancements and new development projects ranged from July 1993 to September 1994, the number of projects engaged in pure maintenance activities increased over the history of the pilot from that shown in Figure 2-1.

Type of Project	Number of Projects
Maintenance	14
Major enhancements (maintenance)	10
Re-engineering	2
New development	5
Total	31

Figure 2-1: Distribution of Projects by Activity (July 1993)

2.2 Project Staff Size

Figure 2-2 shows the variation in project staff size. The largest project consisted of 43 personnel and the smallest was 2. The team staff size includes contractor personnel if used. Five of the projects used contractor personnel in addition to internal personnel.

Number of People on Project	Number of Projects
<5	4
5-9	8
10-24	7
25-49	4
>49	0
No data	8

Figure 2-2: Distribution of Personnel on Projects

2.3 Programming Languages

Figure 2-3 illustrates the primary languages used by the projects. The predominant programming language used by the projects was COBOL; Ada and C/C++ were a distant

second and third. The "other" languages included PC-based Clipper and Dbase. The physical source lines of code (SLOC) varied between 12K and 550K SLOC, but was typically in the range of 35K-40K SLOC.

Programming Language	Number of Projects Using
COBOL	12
Ada	4
FORTRAN	1
C/C++	5
Pascal	1
Misc. assembly	3
Other	5

Figure 2-3: Programming Languages Used

2.4 Computer Platforms

Figure 2-4 illustrates the computer platforms of the systems. The predominant application hardware platform was a mainframe (e.g., IBM and Tandem), but the personal computer and client/server platforms were also represented.

Platform	Number of Projects Using
Mainframe	15
Mini	2
PC/Network	5
Client server	5
Not reported	4

Figure 2-4: Types of Platforms Used

2.5 Tools and Processes

The tools available to each of the projects to capture measurement data varied from excellent to none, depending on the site and project. Several of the sites had excellent project tracking tools. Several of the sites did not have counters for counting lines of code. One site had a separate labor tracking system, but most sites relied on obtaining data from the organization's accounting systems. All of the sites had configuration management tools, but they varied significantly in capability. The effect that such tools had on the ability to collect and use the data is discussed in later sections of this report.

One common characteristic that each site shared was that each had a process improvement initiative underway, and all focused on using the SEI's capability maturity model [Paulk]. Relative to the projects' software processes, it is significant to note that the software maintenance process varied significantly from site to site.

2.6 Pilot Implementation

2.6.1 Overview

DISA/CFSW conducted periodic meetings with the pilot site champions as the primary pilot management mechanism. Initially, the meetings were held monthly. After the fourth meeting, the meetings were held less frequently with two to three months elapsing between meetings. The meeting site was rotated between pilot site locations. The meetings served several purposes:

- To provide a forum for all of the site champions to share their problems and solutions.
- To review the status and assessed progress of each of the sites in establishing measurement processes, and collecting and using the data.
- To provide the site champions with software measurement tutorials and assistance as needed.

The site champions reviewed the status of the pilot measurement efforts at their respective sites, raised issues confronting them, and identified areas requiring assistance. The remainder of the meeting was used to discuss other technical aspects of the pilot, e.g., discussing the pilot database design and reviewing analyses of their data. When appropriate, measurement-related training or general awareness presentations were given (e.g., how to support management with measurement data).

2.6.2 Evolution of the Pilot

The nature of the topics and issues discussed at the meetings evolved over the duration of the pilot. Roughly six different stages evolved:

1. The first stage was dominated by concerns over data definitions, use of the checklists, understanding how and where to get the required data, and negotiating with project managers and support organizations to provide the data. At this point many of the site champions were having their first experience with software measurement while others had limited experience. Given these circumstances, some understandably felt overwhelmed initially. Some sites were collecting portions of the software measurement data prior to the start of the pilot study, and they were able to provide advice on specific issues to those who had little experience in this area.

2. The second stage of issues and concerns focused on data collection forms. Based on experience using the data collection forms and discussions with project managers at the respective sites, the site champions agreed to modify the forms to achieve more consistency between the various forms, clarify formats, and in several cases, collect additional information.

3. Data reporting was the next major concern. The primary concerns to site champions were issues related to timely collection of project data, project manager cooperation, data accuracy, and completeness of data. At this point, it became apparent to DISA/CFSW and the site champions that the procedures used to verify, report, and enter the data into the database needed more attention. It also became apparent that the project managers' cooperation and participation were essential.

4. As sites moved past the data definition and data reporting issues, the design of the database and availability of the reported data became an important issue. The site champions were beginning to look for feedback on the data submitted to DISA/CFSW, as well as guidance for the project managers to use the measurements.

5. After data started to flow back to the site champion, questions and issues relative to the analyses and interpretation of the measurement trends became the key concerns. Relationships between the several core measures provided some insights that were not apparent previously, e.g., effort being charged to completed projects, defect closure rates, or effort applied to enhancements vs. repairs.

6. At the end of the pilot effort, although some sites had dropped out of the program, the remaining sites had plans to continue with the measurement activity. Some sites planned to expand the number of projects involved; still others were in the process

of widening measurement efforts based on their experience with the pilot. In addition, each of the site champions had become a valuable site resource in that they now had the practical experience to establish a process for collecting software measurement data within the realm of the site's software processes.

All sites made progress towards establishing a software measurement process for the projects participating in the pilot. Often, this progress was made in spite of significant organizational stress in the form of organizational restructuring, personnel reductions, and potential base closures threatening or affecting many of the sites.

3. Observations

Much of the discussion in the ensuing sections is an informal assessment of areas that relate to the objectives of the pilot program (see Section 1.2). Many of these issues and observations are likely to exist at other DoD sites and may affect how the DoD (or other large organizations with multiple sites) implements a standard and/or policy requiring software measurement or reporting of software data.

Many observations, issues, and problems encountered during the pilot effort were typical of any project or organization initiating a software measurement program. This was not entirely unexpected, and the literature is rich with lessons learned from such an effort (see [Grady] for example). We have tried to avoid such issues, problems, and observations in this report unless we felt they had a significant bearing on the pilot relative to the objectives.

The reader will note that the observations, issues, and problems discussed under each area may overlap. This is because the effects of the issues and problems frequently have multiple causes. The basis of the information, observations, and issues that are discussed comes from a combination of sources. Thus, much of what is discussed is based on the reports and comments by the site champions. These data were collected from our direct involvement in the pilot. This chapter also includes first-hand observations by the authors and DISA/CFSW personnel. Throughout this chapter, each observation is identified by a bullet and italicized text.

3.1 Collection and Use of the Core Measurements

This section addresses the issues and observations that relate to the pilot's objective to determine the sites' ability to collect and use the core measurements. There were several dominant factors in this area: the measurement definitions, the existence of necessary software processes, the facilities (tools) available to capture the data, and the readiness and ability of the site or project personnel to use the measurement data.

3.1.1 Changes to Measurement Definitions

- *The initial set of core measurement definitions were modified to provide data to better support site project management.*

The application of common measurement data definitions to a diversity of environments and processes was anticipated to be very difficult. Given the DISA/CFSW objectives (Section

1.2) and to the desire to accommodate the differences among sites, the champions originally crafted the definitions with availability of data being the key criterion.

Some months after the original definitions were crafted, it was determined that, with some modifications to the data definitions, the data could provide useful insight into a site's project management processes, in addition to meeting the DISA/CFSW objectives. Figure 3-1 lists some examples of basic project measurement indicators that an organization beginning a measurement program might be expected to support [Baumert] [Rozum92] [Rozum93]. Figure 3-1 then illustrates, using the data definitions, which of those indicators could be applied by a project or site.

As the pilot progressed, an examination of how the data could be used to help the project manager was reviewed by the site champions. The initial measurement definitions required additional detail and some modification to better support indicators that would be useful from a project manager's point-of-view. For example, as initially defined, the data definitions did not differentiate between development and maintenance activity data nor did the definitions separate closed defects by criticality or by type of person who found the original defect (e.g., distinguish closed defects between customer-found defects and those found internally).

Figure 3-1 is a summary of the data collected; the definition checklists and data report forms in the appendices provide a great deal more detail. The columns in Figure 3-1 identify various categories of indicators, as described below:

- *Initial definition:* Identifies those indicators that could be developed with the original definitions.
- Column *A:* Illustrates the indicators that could be supported by the data definitions after they were modified.
- Column *B:* Illustrates the additional indicators that could have been supported by additional modifications to the data modifications.

The modifications to the definitions needed to support column A became the required definitions, while the definitions to support column B were additional information and became optional for the sites. Data from those sites that provided data required in column B were able to be parsed to obtain the level of definition of the data in column A, and were aggregated with the data from sites providing the data required in column A.

Measurement Indicators	Data Collected		
	Initial Definition	A	B
Activity Progress and Status (Planned and Actual)			
# units designed / time period			✔
# units coded or tested / time period			
# units inspected / time period			
# units integrated / time period			✔
# test cases completed / time period			
# successful test cases / time period			✔
Effort (Planned and Actual)			
# total staff-hours/time period	① ②	✔ ②	✔
# staff-hours / process activity			
# staff-hours for rework and changes		✔ ②	✔
Size (LOC)			
# SLOC planned at system test start			✔
# SLOC actual at system test start			
# SLOC in system test / time period			
# SLOC actual at end of system test	✔	✔	✔
Defects			
# total defects/time period			
# total unique defects / time period	✔	✔	✔
# unique defects /finding activity / time period	✔	✔	✔
# open defects/ finding activity	✔	✔	✔
# closed defects/ finding activity		✔	✔
# open defects by age			
# open defects by criticality / time period	✔	✔	✔
# closed defects by criticality / time period		✔	✔
# design changes after units coded			
# code units changed after unit tested			
Schedule-Planned and Actual			
Activities	✔	✔	✔
Work products	✔	✔	✔
Other Indicators			
Project productivity -SLOC per staff-hour	① ②	✔ ②	✔
Development defect density - # defects per SLOC	✔ ②	✔ ②	✔
Rework to total staff-hour ratios	① ②	✔ ②	✔

Legend:

	Indicates no data collected
①	Definitions do not separately identify development data from maintenance data
②	Definitions do not address planning data
✔	Definitions address this data

Figure 3-1: Measurement Definition Comparisons

The primary difference between columns A and B in Figure 3-1 is a project's ability to separate maintenance data from development data and to separate collection of defects closed by finding activity (e.g., testing vs. customer discovery) and criticality. The site champions also agreed to attempt (but were not required) to collect the data needed under column B which includes the collection of planned as well as actual data for size, progress, and effort. The revised definitions (column A) allowed the project manager to distinguish and assess the status and progress of the development and the maintenance activity, and separate problems found by customers from those found during testing.

While it did not have a dramatic impact on the pilot, modifying the data definitions did have a ripple effect on the pilot effort in two areas. First, the data collection and report forms were modified to accommodate the new definitions. This, in turn, required that each of the site champions review the revised forms with the people collecting the project data and then change their collection process to collect data according to the new definitions. The second effect from the changes was on the pilot database schema and data entry screens. The schema and screens were based on the data report format and also had to be redesigned to comply with the new definitions.

The chances of encountering this type of change in developing a measurement program can be significantly reduced by ensuring that the measurement data definitions can be used to address the goals and issues of the organization collecting the data before multiple sites and projects implement a data collection process based on the measurement definitions.

3.1.2 Data Collection Forms

- *The design and content of the data collection forms should be tailored to avoid confusion, inaccuracies, incompleteness, and misinterpretation of reported data.*

At the initial training session, the pilot participants decided to use the reporting forms (also referred to as data collection forms) in the respective SEI technical reports ([Florac], [Park], [Goethert]) as a starting point for their own reporting forms. Some of the forms suggested in the SEI technical reports were modified for use in the pilot, while others were designed during the pilot workshop.

As the site champions gained experience using the forms, they reported that many of the initial data collection forms were incomplete and lacked sufficient direction for use by the data collectors. In addition, several of the forms did not provide for sufficient supportive data, and in one case, led one to believe a great deal more data were required than were actually needed. The forms also provided for entry of summary data as well as the detail data, but the summary data were not identified as such. The site champions also concluded, during a review of the forms at an early site champion meeting, that they, as a group, were not certain how to interpret several of the forms.

Eventually, the SEI reporting forms in the SEI technical reports, although discussed in those reports as examples, were significantly tailored by the pilot measurement program. Several of the site champions took on the responsibility of modifying the forms to eliminate many of the problems that were being encountered. At the same time, the forms were modified to reflect the changes to the defined measurements.

Another problem was the lack of a written guide (process) instructing sites and champions how to use the reporting forms. This later appeared to contribute to problems and misinterpretations in data reporting. These problems further led to confusion during analysis of the data contained in the database. The champions were strongly encouraged to develop a data collection and reporting guide specific to their site.

- *Data verification (consistency and range checking) and audit control of reported data are essential to avoid erroneous data in the measurement database and gaps in the reported data.*

The initial database mechanism provided no data verification or editing capabilities. The site champions were asked to review the data collection forms each month before they were submitted for entry into the database. The champions were to check that the forms were completed correctly and completely, to check for consistency and reasonableness, and to catch errors in arithmetic and spelling. The data collection forms were delivered to DISA/CFSW as hard copies and were, in turn, copied and delivered to the contractor responsible for the database. The contractor then updated the database, but the initial database mechanism provided no data verification or editing capabilities. Additionally, two separate organizations were executing the procedures used to ensure that the completed data collection forms were archived and controlled while, or after, the data were entered into the database. The procedures were not well defined, leading to configuration management problems and missing data in the DISA/CFSW database and archives.

The above procedure was used for the first several months of data collection. Errors and inconsistencies in the data were not noted until DISA/CFSW staff and advisors tried to conduct preliminary analyses of the data for the sites. During efforts to correct the data, the need for improved archiving and control of the reported data was requested by the site champions and implemented by the database contractor and DISA/CFSW staff. Nevertheless, it was necessary to request the site champions to review the apparent inconsistencies and correct the erroneous data where possible.

The inability to conduct data verification and audit control during the initial several months of the pilot resulted in extra effort on the part of the site champions and delayed the analysis of the reported data and feedback to the participating projects. These types of problems can be avoided in start-up efforts, to the extent that the process can be predefined, by establishing a rigorous data management process at the start.

3.1.3 Changes to Software Processes Required to Collect Data

- *Consistent collection of the SEI core measures depends upon the existence of several software processes consistent with several level 2 key process areas (KPAs) in the SEI capability maturity model (CMM).*

Several of the sites were collecting software measurement data prior to the start of the pilot effort. The existing site software processes at these sites, however, typically did not include collection of all the project or process data required by the pilot software measurement definitions. Collecting all of the measures defined by the pilot required at least one major process change at most sites. The measures defined are those that might be considered typical project management data. As such, the processes that produce, collect, and use the data are also consistent with several *activities to perform* within level 2 KPAs of the CMM.

To collect all or some of the measures, a project may have had to make changes to typical CMM level 2 processes, and sometimes implement an entirely new process. For example, to collect defect data, several projects had to initiate a process for tracking defect reports. Other projects had to establish processes that included the reporting of staff-hours for one or more of the different types of employees (e.g., contractor employees, government civilian employees, and military employees). Another example is projects modifying their project scheduling and tracking process to provide the schedule measurement data required by the pilot.

Efforts to introduce a measurement activity into the existing software process initially delayed data collection for some projects, and in a few cases, a project never succeeded in establishing a software process that would yield a complete set of measurement data as defined by the pilot. The status and capabilities, or maturity level, of the existing software processes at the sites were an important factor in determining the effort and time to install a measurement process.

Note, however, that even mature software processes would probably need to be modified to accommodate the collection of software measurement data to a definition developed by an entity external to the project.

3.1.4 Availability and Capability of Data Capturing Tools

- *The availability and capability of tools used to capture the data affect an organization's ability to collect defined measurement data.*

Each of the pilot sites was encouraged to use existing facilities to capture the defined measurements (e.g., labor tracking systems, configuration control systems, problem tracking

systems, and project management systems). Many of the data capturing systems were not designed to capture software measurement data as defined by the pilot. As might be expected, every site initially had difficulty collecting the data for one or more of the defined core measures because of the availability or capability of the data capturing tools at the site.

For example, to obtain effort data a few of the sites had access to and used labor accounting systems or project tracking systems that were functionally capable of providing much of the effort data required by the measurement definitions. Typically, the systems were designed to be used by software organizations and provided much of the detail pertinent to software effort analysis. On the other hand, some sites had to rely on labor accounting systems that were not designed to capture certain attributes of software effort measures. For example, some were unable to obtain work breakdown structure data that detailed the staff-hours expended for design, coding, and testing. Some systems were unable to capture contractor effort, and others were unable to capture effort of military personnel. In several cases where the data were extracted from administrative reports, the data were not timely relative to the project managers' needs, and did not support their efforts to balance workloads or control budgets.

Other projects initially could not capture source lines of code counts because they did not have a code counting program. In one case, this was because the project was a legacy application that ran on an old hardware platform, and resources were not available to create a code counter. More frequently, a project had not yet obtained or used a code counter prior to the pilot and was dependent on others to provide a counter and instruct the project how to use it. Sites had hoped to be given a single counting tool, but due to the wide differences in languages, platforms, and application architectures, a single code counting tool was not possible.

Some sites had configuration management systems to track change requests to the software, but the system did not differentiate defects from system problems or enhancements; thus, site personnel had to parse the change requests manually to track defects. Some projects had problem tracking systems that tracked at the problem-report level and did not differentiate between defects and other problems associated with the system. For several of the required data collection fields, each problem report had to be examined manually to determine its appropriate category. For example, the number of defects discovered by the software user and those uncovered by the programmer testing the software had to be determined manually by examining each problem report.

Several sites had project management systems that tracked the various projects' work efforts and commitments in terms of schedules, work item size, responsible parties, and status. These systems were useful sources for much of the measurement data, but frequently required manual extraction of the data in lieu of a periodic report.

With some exceptions, the sites were able to use existing data capturing tools as a source for the measurement data, or obtain such tools; however, it frequently required some manual intervention or manipulation to capture the measurement data as defined by the pilot. In addition, success in using the existing data capturing tools was due to the ingenuity and resourcefulness of the site champions.

3.1.5 Data Collection Process

- *Documentation of the data collection process helps an organization to understand and implement the process and then to improve and sustain the process.*

The pilot process did not include a process guide documenting the measurement definitions and how the site champions were to report the data to DISA/CFSW. Such a guide was later used on another DISA/CFSW effort to collect data over multiple sites, with positive results.

A process guide developed by the organization sponsoring the collection of data from different sites would have helped to establish the collection process and communicate more clearly (than the SEI checklists alone) who had the responsibilities to provide what information.

Each site that documented its own data collection process found that having the process documented was useful. However, a general process guide for DISA/CFSW on the expectations and responsibilities of the sites would have addressed several questions that arose when the sites documented their processes and responsibilities of the sites. Based on comments from several site champions, the documented data collection process for a site resulted in several benefits:

- Provided a clear definition for each of the data items that were to be counted on the data collection and reporting forms. The process guide gave instructions as to who should collect the data, when and where it should be collected, and who should receive the completed form. Examples of the data collection forms were provided, in addition to copies of the checklists defining the data to be collected.
- Related the collection process to the sites' software process and supported a site's overall process improvement effort.
- Became an instrument to inform other site employees of how the measurement process worked and how the measurement data would be used. Describing how measurement would be used also reduced some of the resistance to measurement by managers at the site.
- Provided continuity in the data collection process when champions changed. Those sites that had a documented data collection process had fewer problems transferring

the measurement task to the new champion. The documented process helped replacement champions become familiar with the measurement process quickly.

3.1.6 Participation of Project and Site Management

- *Successful software measurement programs require the project manager's participation and involvement in the measurement process.*

Project managers must understand the purpose of the measurements and become confident that the data will not be used against them or their project. Several site champions reported that, even though the site commander supported the pilot effort, they experienced resistance from the project managers to provide the data. It frequently took several months for the site champion to reach agreement with a project manger to provide data.

There appeared to be several reasons for this resistance. Based on the reports and comments of the site champions, many of the project managers (and others) did not appreciate the purpose or objectives of the pilot activity or how the data collected would help their project, and they looked upon the activity as a nonproductive exercise with which they did not have time to comply. In addition, the project managers were fully aware that the collected data were to be reported to DISA/CFSW and stored in a DISA/CFSW-managed database, and it was reported that some of them were concerned how that data might be used. The site champions made it clear that informing the project managers that the pilot was part of a DoD pilot effort and that they could use the data to improve their software project processes, was not sufficient reason to expect the cooperation of project managers.

Because they did not know how the data were being used, a few project managers, while agreeing to cooperate, "protected" themselves by reporting only the data they wanted to release. For example, one project would determine which and how many defects were reported to DISA/CFSW in the monthly reporting of defects. Also, some projects did not report all defects discovered. Those that could be corrected quickly and simply were never logged; thus, they reported only the "important" defects.

Project managers were more responsible and cooperative in those instances when the site champion went beyond just collecting the data. It appears that a site champion must be able to demonstrate to the project manager that he/she and the data can help the project. Those that were successful in doing this typically analyzed the project's data and reviewed their interpretations of the information with the project manager.

This experience re-emphasized the need to create and use a project-related measurement plan for implementing software measurement. The project goals, the measures, the data to be collected, the processes for collecting the data, and the analysis and reporting of the

measurements related to the project need to be identified and defined early in the pilot with the involvement and participation of the project manager.

- *Senior management oversight reviews of projects using the measurement data are an important factor in getting measurement into practice in an organization.*

All of the sites' senior management were supportive of software process improvements and measurements. Those site managers who conducted periodic reviews of the projects participating in the pilot were key factors in successfully establishing a software measurement process at their site. Oversight reviews are one of the key process activities for CMM level 2 and are an essential ingredient for software process improvement. Software measurements provide the quantified information relative to project progress, issues, and problems, and therefore, are an essential ingredient of oversight reviews. In this context, the oversight review gives reason and purpose to software measurement.

To make the importance of management oversight more visible to senior management, the topic of oversight reviews was brought to the attention of the site executives during a conference held at the SEI for executives of the pilot sites. As a result, several of the sites instituted senior management reviews with the participation of the site champions. Sites that did conduct periodic oversight reviews during the pilot were successful in collecting data, having project management use measurement information, and expanding the measurement process to include additional projects at the site.

3.1.7 Education/Training

- *Software measurement training is needed for software process personnel to plan, implement, and support a software measurement process.*

Before the pilot began, few of the site champions had been trained or were experienced in establishing a software measurement plan. A four-day training session introducing software measurement and its uses was provided by DISA/CFSW at the outset of the pilot, which included training on using the SEI checklists. As a result, the site champions had an initial exposure to software measurement and could interpret the definition checklists.

As the pilots progressed, DISA/CFSW recognized that the initial training was insufficient and that most of the champions could benefit from additional training. For site champions, training in how to develop and establish a measurement process would have been useful (i.e., training in how to establish goals toward which measurement is applied, define measurements, develop data collection procedures, collect the data, and analyze or use the data).

Each meeting with the champions generally had a training component as part of the agenda. The champions were given various types of training covering data collection procedures, using the measurements to make decisions, establishing goals and measurements to measure progress, defining data, and using software measurements to establish trends and detect problems.

A one-day workshop on establishing goals, and identifying measures to use in support of the goals, was piloted for projects at one of the sites. Due to resource and schedule constraints, however, it was not presented to the other sites.

Many of the champions indicated that the tutorials would have been very useful before starting the pilot, as it would have made their tasks less difficult. Assuming this is the norm throughout the DoD sites, there is a need for future site champions and/or software engineering process group (SEPG) members to be trained and given guidance in developing a software measurement implementation plan that supports the software process improvement initiatives underway at the various DoD sites.

- *Project managers and site executives need assistance on how to use the software measurement data in their decision-making process relative to their projects and business objectives.*

The site champions consistently raised the issue that they and their project managers needed assistance on how to use the software measurement data that they had collected. Many of the project managers did not have experience using objective data in making decisions and required assistance. Based on the site champions' requests, several tutorial sessions were conducted for the site champions, with the intent that they would then be able to assist the project managers in using the data. Often this did not work out as planned because project managers would ask questions that the newly trained site champions could not address. Providing managers some training, though, may have helped facilitate the champion and manager working together as a team to use the data.

3.2 Applicability of Common Definitions

One of the pilot program's objectives was to determine the extent to which common definitions could be used across different applications, domains, and organizations. The reason for common definitions is to facilitate communication among the several programming organizations and to form the basis for some level of comparative analysis, e.g., to improve the estimation process. This section discusses several of the issues that the pilot identified relative to the use of common definitions of the measurements. Primary issues that affected the extent to which common definitions could be used included existing software and

administrative processes, pre-existing software measurement data, and the context of measurement data.

- *To use commonly defined measures, the approach to collecting the data needs to be integrated with other software and business related activities .*

Some sites had activities in place prior to the pilot to collect and report on software measures other than those defined for the DISA/CFSW pilot. This data collection was generally done independently in different parts of the organization. When this was the case, it was not always understood what the data currently being collected represented or how the required data were different from that already being collected.

Collection of the defined data usually required some sort of modification to the existing software process. Before changes to the process could be implemented, the site champions first had to determine the definition for the data currently collected, how the common definition of the DISA/CFSW pilot compared to the site's current definition, and then, if and how a site process would need to be modified to collect the data according to the DISA/CFSW pilot definition. Modifying the existing software process sometimes affected different parts of the organization. That is, a site champion would have to talk and negotiate with different parts of the organization where the data were already being collected. Having the champion be the point-of-contact between the processes and the person integrating the collection of data was indispensable in a multisite approach such as the pilot.

While this effort by the champions was unanticipated, it paid off by establishing a more cohesive data collection process and resulted in an improved understanding of the software processes and the data being collected.

Some sites collected a measure different than a core measure for seemingly the same purpose that a core measure was to provide. In one such case, a pilot site was collecting function point date in lieu of SLOC. That site was strongly encouraged to provide both the function point data they collected as well as the SLOC data for the core measure.

- *Differences between an organization's existing software processes and administrative procedures can lead to misinterpretations of what are thought to be commonly understood software measurement definitions.*

Each item within an attribute category of the checklists needs to be clarified with regard to an organization's software process. For example, differences in the software maintenance process among the sites were a significant factor affecting the ability to use several of the measurements as a common reference point. There seemed to be at least three different ways in which maintained (repaired and enhanced) software was released for use. Some projects released an enhancement or repair action as soon as it was written and tested (this

process was used primarily by projects supporting an installation operating on the site); such a process could theoretically have a release every day. Other projects released whatever software was completed periodically (e.g., all repairs and enhancements ready for release each quarter). Still others released the software according to a schedule based on the functional content or effort required to implement the software changes.

The different release processes, in turn, often dictated different software processes leading up to the release of the software, e.g., program change management, configuration control management, problem management, design, coding, and test, as well as the way in which the software measurement data were collected and counted.

Typically, software measures, such as schedule and progress, have different meanings in each of the maintenance processes outlined above (as it did for the pilot sites). The common schedule definition (assigning completion dates to milestones and deliverables) was used differently by those using the project maintenance processes outlined above. In these cases, completion dates were provided that satisfied the definition criteria, but since the data had a different context than another site's data, it was prone to misinterpretation in the absence of knowledge about the context of the data. In this case there was an inadvertent implied assumption made about a common maintenance process when defining the schedule measurement that would not be apparent until comparative analysis was attempted.

The measurement of software size caused problems for several sites because it was defined to be the total number of non-commented source lines of code as measured at the completion of system test. Several projects released "bug fixes" every day, while others released system updates every three months, with each release undergoing a system test. The monthly software measurement data were collected every month. It became clear that size data in these cases did not represent the same thing from month to month or from project to project.

Some of the pilot sites had problems dealing with the common definitions because the reporting frequency of the software measurements was different from the frequency of existing administrative process cycles. For example, some sites had biweekly accounting periods, some twice per month, and others monthly, causing differences in the collection and counting of effort (staff-hour) data between projects and between sites.

Whenever obstacles like those described above arose, a work-around was developed when possible, or sometimes the measurement definitions were changed to accommodate those with the obstacle, thus making the collected data relevant to the collecting project. However, the inter project communication and comparison issues related to the common definitions were never completely resolved because of differences in the software processes and administrative procedures among the sites. These differences made it difficult to support site requests from the champions to aggregate data or understand

differences when comparing projects to other sources (e.g., the commercial databases provided by software measurement vendors).

3.3 Cost of a Measurement Program

Data were collected at the sites on the amount of time (staff-hours) that was expended by the site champions and others in the organization to implement the collection and reporting of the SEI core measures. The following sections include charts that show the staff-hours used by each site to support the pilot, and the distribution of effort by all the sites over the first 13 months of the pilot. The data are not to be considered averages or expected ranges, but merely what was observed for this pilot. There are more than likely forces not captured by the data that are constraining the amount of effort each site expended towards the measurement pilot. From the view of trying to understand the cost of implementing a multisite software measurement initiative, the data collected do seem to support the following 3 points:

- The amount of site support required is largely dependent on the site's software infrastructure.
- The site's support effort is only moderately increased with more projects.
- The total site support requirements tend to decrease as the measurement program matures.

3.3.1 Pilot Site Effort per Site

- *Wide variation in the number of staff-hours expended by a site collecting software measurement data should be expected.*

The total staff-hours spent by each site is shown in Figure 3-2. Two bars are shown for each site. The black bar indicates the total number of staff-hours spent by all personnel at that site to support the pilot (including the site champion effort). The gray bar indicates the total number of staff-hours spent solely by that site's champion.

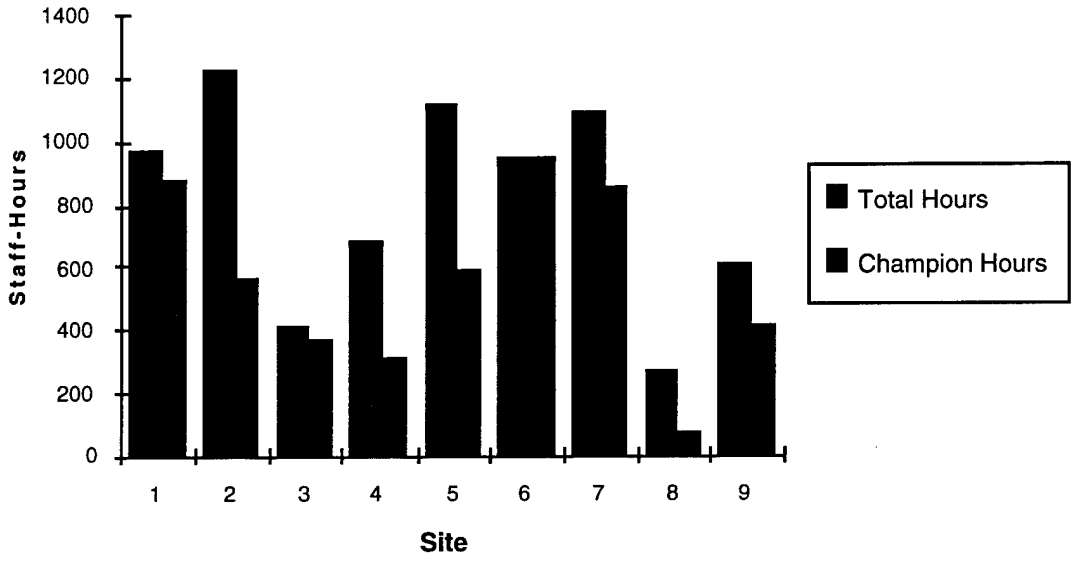


Figure 3-2: Pilot Site Effort to Support Measurement

The average number of total staff-hours spent by the sites was 814 with a standard deviation of 315 hours. The minimum total number of hours spent by a site was 269, and the maximum total hours spent by a site was 1225.

The average number of staff-hours spent by a site champion was 555 and the standard deviation was 278. The minimum number of hours spent by a site's champion was 72 and the maximum total hours spent by a site's champion was 947.

There are many factors causing the above variation in the amount of effort spent on software measurements for the sites. Based on our observations, the following issues are all key factors in the amount of effort required of the site personnel, and they varied from site to site:

- Measurement processes that were already in place and the degree to which they could be utilized.
- Software processes that were in place and could be used for leverage.
- Experience defining processes that could be used to define the measurement process.
- Amount of tailoring of the measures that was needed relative to the organization's practices.
- Resistance inside the organization.

- Number of projects participating.
- Experience of the site champion.

3.3.2 Overall Cost of a Measurement Program

- *The cost of the overall program is not proportional to the number of projects participating.*

There did not appear to be any single factor that could explain the variation between the sites. In particular, there was a very low correlation value (.01) between the number of projects participating in the pilot per site and the total number of staff-hours expended by the site.[4]

Because the analysis above did not show a significant correlation, we then looked at the correlation value between the ratio of champion staff-hours to total staff-hours and the number of projects [i.e., (champion hours/total hours) / number of projects] at each site. A correlation value of -0.58 was obtained which indicates that the number of staff-hours required for additional support at the site compared to the effort expended by the site champion tends to increase (i.e., the ratio decreases because the denominator is increasing) with the number of projects involved in the measurement activity.

We believe that part of this phenomena is due to the constraint on the number of available champion staff-hours. This more than likely constrained the number of champion hours expended, but also constrained the number of projects that participated, i.e., the number of projects that a single champion could work with. There did appear to be other factors affecting the cost of implementing the collection of software measurement and data collections. Factors that we have identified are related to the capability or readiness of the site's infrastructure, e.g., training, processes, tools, experience. A person responsible for implementing a multi-site measurement program should understand that the number of projects involved in the effort probably only moderately influences the overall cost of the program.

---

[4] The correlation value measures the relationship one variable has on another variable. The value can range from negative one (-1) to positive one (+1). A value at or near zero implies that there is little or no relationship between the two variables, i.e., if variable A increases, variable B may or may not increase, and it's movement is independent of A. A correlation value that is less than zero implies a negative or opposite relationship, i.e., if variable A increases, variable B decreases. A correlation value greater than zero implies there is a direct relationship between the two variables and both move in the same direction, i.e., if variable A increases, variable B also increases.

3.3.3 Pilot Site Effort per Month

- *The average total effort decreases over the first year of the effort.*

The graph shown in Figure 3-3 shows the total and average staff-hours for all sites distributed over a 13-month period (July 1993 through July 1994). The graph shows a decided decrease in the number of staff-hours devoted to the pilot with the passage of each month. The staff-hours per month reflect the hours reported by the reporting sites. The average site staff-hours per month reflect the average of the sites reporting data. The average monthly staff-hours appears to be following the trend established by the previous 12 months, i.e., a gradual decrease from 85 staff-hours per month to about 50 staff-hours per month at the end of 13 months.

This decrease can be partially attributed to the reduction in site champion meetings from once per month to once every three months. Additionally, part of the decrease can be attributed to the sites becoming more proficient in collecting and reporting the measurement data.

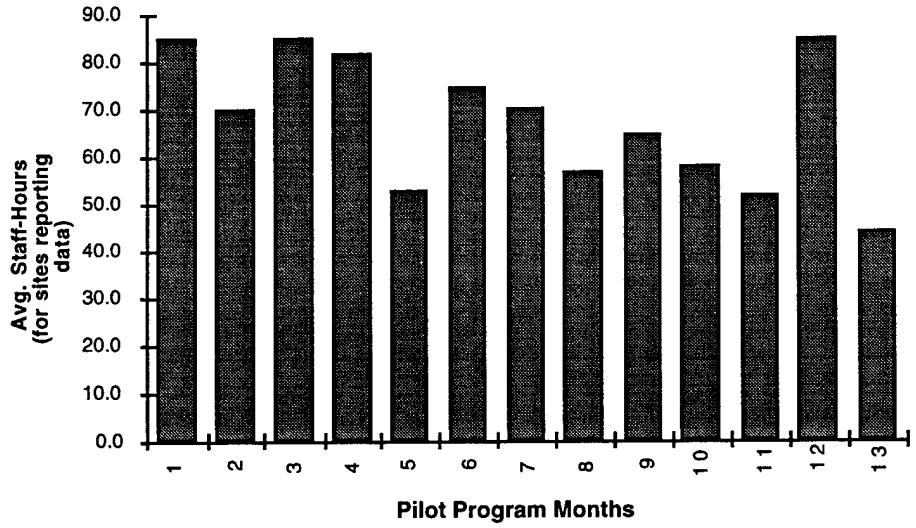


Figure 3-3: Average Monthly Effort Spent by Pilot Sites

3.4 Analysis and Reporting of Measurement Information

In this section, we discuss the observations and experiences encountered during the pilot program relative to the analysis and reporting of measurement information at both the organization and corporate level. Four perspectives are used to discuss the analysis and reporting of measurement information:

- A project's and organization's use of its measurements to track status and progress.
- Project use of aggregated project measurements as benchmarks.
- Headquarters (corporate) use of specific project measurements for comparative analysis.
- Headquarters use of aggregate project data and measurements to determine overall trends and status.

3.4.1 Getting Projects to Use Data for Project Management

- *Progress in getting projects to use data from the measurement program in project management activities will vary significantly.*

After the first 9 months of the software measurement pilot, the champions from 3 sites (representing 10 projects) reported that the projects were regularly using the measurements to track status and progress, as well as to identify issues and problems encountered in the software process. One of the 3 sites used the measurements as part of regular management oversight reviews, e.g., in-process reviews, project plan approvals, and problem control management. The site commanders at these 3 sites had directed, or were in the process of directing, that the measurements be extended to all projects at the site.

The remaining six sites were collecting measurement data, but had not yet reported that the data were being used by the projects to track or manage status or issues. The champions from several of these sites reported that they still were having difficulty collecting the data, while others reported that the project managers continued to ask questions regarding the utility and application of the measurements.

At the same time, DISA/CFSW outlined a plan to provide feedback and analysis of the reported measurements to the sites. The plan consisted of providing the sites

- A copy of the pilot repository with the respective site data.
- Prompt feedback confirming data had been received or concerns about suspected data anomalies.
- Graphical reports with supporting analysis regarding that site's data.

The purpose of the graphical reports and analysis feedback from DISA/CFSW was to stimulate further analysis by the projects and to illustrate the way in which the measurements could be used.

Over the ensuing three months, the DISA/CFSW staff conducted an intense analysis of each project's data, identifying anomalies and raising questions and issues suggested by the measurement trends. Some project data were not complete or sufficiently accurate to allow analysis, while other projects had adequate information to allow trend analysis and comparisons with related measurements within the project. The DISA/CFSW staff and

advisors were not sufficiently familiar with all the projects' processes and working environments, and therefore could not make a conclusive analysis; they were, however, able to raise questions and issues that may be of concern to the project manager. The site champions felt that this feedback was instructive and useful in demonstrating how the measurements could be used by the project managers. Those that had reviewed the analysis with their project managers also commented that the managers found the information useful.

3.4.2 Project Benchmarking With Measurement Data

- *Motivate organizations by providing benchmarking results.*

DISA/CFSW initiated a study that would provide the sites additional feedback on how they were positioned relative to other software projects in government and industry engaged in similar types of software environments. The PADS (Productivity Analysis Database System) created and populated by QSM (Quantitative Software Management, Inc.) was used as the basis for this benchmarking activity. Only those sites that volunteered, and were able to obtain some additional data, were involved in this study. The study positioned each of the participating sites' projects on a productivity index and a staffing index based on the size and category (domain) of their projects. The site champions' response to this study was very positive and they indicated they were eager to share the results with their project and site managers. The site champions indicated that such data would be an excellent motivation, regardless of the project' s position on the index relative to the norm.

Given the favorable reaction by the site champions to the benchmarking analysis, it would appear that the establishment of a corporate-level repository that could be used to conduct such analyses would facilitate, encourage, and motivate projects to use software measurements.

3.4.3 Comparative Analysis

- *Corporate or headquarters measurement programs wishing to make comparative analyses of projects need to collect sufficient contextual data or characterization information on projects.*

As indicated in earlier chapters of this report, DISA/CFSW originally had no specific intent to analyze the reported data, other than to assist the sites and projects in the software measurement process. To encourage and motivate projects at sites, later in the pilot, DISA/CFSW staff and its advisors helped champions to analyze project-specific data, but only to help the projects use the data to manage their projects.

There was no attempt by the DISA/CFSW staff or advisors to conduct analyses comparing the defect rates, productivity factors, schedule performance, effort expenditure, etc., across the participating projects. As discussed earlier, the projects used different processes, languages, and supporting tools; and supported a wide variety of defense department functions. Therefore, the measurement data often had a different context from one project to another. In fact, it is very difficult to collect information characterizing these differences, and any attempt to collect the information at such a level of detail as needed to compare projects was out of the scope of the pilot. Any comparison of the projects without this information to normalize the data would be very misleading. Comparisons that turned out not to be misleading, in our experience, have been pure coincidence.

3.4.4 Corporate Profile and Process Status and Trends

- *The data, with appropriate contextual information, can be used to report the current status and establish trends in an organization's software process.*

The enterprise or corporate interest in software information is presumed to be analogous to that of a corporation stockholder (e.g., the balance sheet, profit and loss statement, and cash flow analysis). The difference, in the case of software, is that the corporate interest is in terms of project profile and process trends and status (e.g., languages, technology, tools used, product inventory, overall quality factors, customer satisfaction, skill levels and needs, and overall budgets and costs).

Much of this information may be derived by periodic collection of project profile data and certain quantitative summaries of project size, effort, defect, and schedule measurements.

An example of the kind of information that can be derived from such project profile data is given in Chapter 2 of this report. All of the information cited in Chapter 2 was recovered from the portion of the DISA/CFSW repository containing profile information about each project. The repository established by DISA/CFSW served dual purposes. The first was as a convenience to the participating projects. The repository provided a database for the projects to store their data without having to develop their own. It also served to store information and data necessary for the pilot program, namely the project profiles and the effort measurements associated with the pilot program. Thus, the repository served both the project level and the corporate level, albeit, at a relatively basic level.

The profile data collected for the pilot program was useful, but, viewing the data in hindsight, there are many more pieces of information that a corporate-level information repository might need. For example, contextual information characterizing the projects' processes, tools, skills, and application domain needs to be collected and structured in the repository so that the data may be readily accessed using data analysis tools. Certain other data, related to the goals and issues at the corporate level, need to be collected, aggregated, and related to the project profile data. At the corporate level, very little of the detail on the measurement data that is collected at the project level, week in and week out,

is needed or even desired. By the same token, the project manager has little need for the aggregate level of data used at the corporate level for tracking projects' status and progress.

3.5 SEI Core Measurement Checklists

- *The SEI checklists were helpful, but improvements and revisions are needed.*

The SEI checklists were used throughout the pilot and facilitated communication among the champions from the sites, as intended. At the initial training session, DISA/CFSW used the checklists effectively to describe the measures that would be collected and then used the checklists as the means to define the measurement data that would be collected. Several times throughout the pilot, the checklists were referred back to in order to address the intent of various definitions of the data that the sites were to collect. When measurement definitions were modified, the checklists were used to guide the process and, in this way, were a valuable part of the pilot.

While using the checklists rigorously, several issues and problems were noted. For example, inconsistent and differences in form and format across all of the checklists meant that the people using the checklists had to reset their thinking on how to use that particular checklist and use a different process each time another checklist was used. Examples given were the need for consistent information in the form header about the project or product, common formats to describe array data, use of the same terminology to describe the measurement attributes, and similar overall structure and architecture.

Other issues dealt with the individual checklists. For example, it was pointed out that while the checklists supported software development activities, modifications were required to address software maintenance activities adequately.

Problems arose when trying to use the example data collection and report forms provided in the SEI checklist reports. Initially the site champions tried to use modified versions of these example forms. Within the first three to four months, many of the site champions realized that the data collection forms needed revision to make them more understandable and usable. In addition, the site champions realized that changes to the forms were necessary to differentiate between maintenance and development activities and provide more information about problem reports and status. Ultimately, a new set of data collection forms were developed that were tailored directly to the data collection needs of the pilot program.

To help identify and resolve problems with the checklists, document comment forms (DCFs) were written for each issue or problem noted during the pilot. Copies of the DCFs are contained in Appendix A.

3.6 Measurement Program Guidelines

The several observations concerning measurement program guidelines that are offered in this section are based on the experiences, observations, reports, and comments of all the pilot program participants, DISA/CFSW staff, advisors, and site champions.

3.6.1 Organizational Goals

- *For a measurement process to be successful, it is essential to establish goals or objectives that map a sense of purpose to the measurements for all of the participants in the measurement process.*

The DISA/CFSW pilot measurement program had its own objectives (see Section 1.2), and the pilot sites worked towards collecting data relative to those objectives, essentially fulfilling their commitment to participate in the pilot. When the DISA/CFSW pilot objectives are examined, it is clear that the overall goal was to assess the cost and ability of a population of diverse organizations to collect the SEI core software measurements relative to a common definition. In addition, DISA/CFSW wanted to assess the feasibility of using common measures while, at the same time, having the sites use the data to improve their software processes. However, while these goals served DISA/CFSW's purposes, they did not necessarily serve the purposes of the sites or projects. Some of the projects and sites did not view their organizations as stakeholders in the data collection and use of the pilot measurements as evidenced by the reports of the site champions and the reluctant participation of some project managers at the beginning of the pilot program.

In establishing a multisite, multiproject measurement process, it should be remembered that various organizational entities (e.g., corporate, division, program, staff, and project) often have different specific goals, issues, perspectives, and interests, even if the overall goal is harmonious among the entities. Often this results in different measurement needs and project priorities; and to the extent these organizational entities are participants in the measurement process, their measurement needs must be addressed.

3.6.2 Site Champions

The background and experience of the site champions participating in the pilot program were wide ranging. Several were members of the site SEPG group, one was an active project manager, and several were very familiar with the issues relative to process management and software measurement; however, not all were familiar with the issues that a project manager might face in managing a project. Several of the champions were very proactive. Others needed a significant amount of assistance and advice in establishing a data collection effort, working with project managers, and knowing how to use the software

measurements. Nearly all had additional responsibilities. Without exception, all champions were enthusiastic, could communicate effectively, and had their site executives' support.

- *The institutionalization and acceptance of a software measurement program is facilitated by site champions who serve as technical change agents for the organization or projects and act as the liaison to a headquarters or corporate software measurement staff.*

The DISA/CFSW pilot program was heavily dependent on the site champions to initiate and implement a measurement process for each of the participating projects. These site champions also served as a focal point to define and communicate the software measurements at both a project level and with other sites. The site champions were familiar with the project managers and had the opportunity to communicate and work with them on a day-to-day basis, a task that the DISA/CFSW staff could not have accomplished nearly as effectively. Having a local focal point for software measurement who is familiar with the site's projects, personnel, and software process proved to be very advantageous, and the pilot probably could not have been successful without them.

- *Establish or provide criteria for the selection of site champions to achieve the best results.*

Based on the reports and comments made by the site champions themselves, several characteristics seemed to be important differentiating factors in successfully initiating and making significant progress towards establishing a measurement process. Two important factors that seemed to relate to the site champion's success were credibility (by both the project mangers and site managers) and a proactive approach. Those champions that were considered to be peers by the project managers by virtue of their experience or reputation were able to quickly establish a good working relationship with the project managers. Similarly, those that were proactive in their relationship with the project managers, showing them how the measurements could be of help and helping them to use the measurement data, were also able to make quicker progress in establishing a measurement process.

- *Expect and plan for turnover of the personnel assigned as site champions.*

Turnover of the site champions caused delays in the establishment of a measurement process for several of the projects in the pilot program. At the first working group meeting after the initial training, seven sites were represented by site champions. Within the first six months of the pilot study, four of those seven sites had new champions assigned. In the interim, two sites were added and the new site champions were trained. Of these two sites, one champion was changed within six months. During the last six months of the pilot, two sites assigned new champions.

The turnover meant that the new site champion had to be trained on the SEI measurement checklists and then on the pilot's measurement process and definitions being used. Often,

additional delays occurred when the detailed process of collecting the site data resided cognitively with the former site champion. This implied that new site champions had, at a minimum, two areas and processes to learn: (1) the measurement definitions and process being used for the pilot, and (2) the specific data collection process being used at the site. Additionally, sometimes the site champions had little experience with software measurement and would require additional instruction in basic measurement principles.

To help with the turnover problem, the DISA/CFSW staff and advisors recommended early in the pilot program that each site document its process for collecting and reporting the measurement data. This was an action item taken by each site. The several sites that had better measurement process documents suffered less impact due to site champion turnover and had fewer problems in implementing and sustaining the software measurement effort.

3.6.3 Support

- *A trained staff to provide analysis, feedback, and support will sustain the implementation and the collection of software measurement data.*

As the pilot progressed, the DISA/CFSW staff realized that site champions and the project managers needed assistance in understanding how to use the data being collected to help manage their projects or improve the process. To help with this, presentations and numerous reports were given to the site champion at the meetings to familiarize the champions with project management measurements and how they might be used. Additionally, a one-day workshop was piloted at one site which included the project managers and key technical personnel. The workshop was intended to illustrate how project managers could use data to support their management processes and address issues and problems specific to them.

In the later stages of the pilot effort, the reported data were analyzed by the DISA/CFSW staff and advisors from the point of view of the project manager, resulting in a series of questions about the project status and progress that the project manager might investigate. The site champions reported this was quite useful to both them and the respective project managers in that it gave them several new and different perspectives to consider that were useful in improving the project's performance and process, thus reinforcing the notion of establishing a software measurement process.

4. Lessons Learned

Although the lessons learned that follow are by no means all inclusive, they are what we believe need to be considered by an organization when planning for the implementation of a measurement program across multiple sites or projects, or by the DoD when it considers a DoD policy on software measurement. As in Chapter 3, we discuss the lessons learned in the context of the pilot program objectives with a focus on lessons that relate to multiple sites, domains, or application types.

All of the lessons learned are based on the experiences and observations gained from conducting the DISA/CFSW pilot program described in the previous chapters, as well as our experiences in establishing measurement programs in other organizations. Some of the lessons learned are based on activities that the pilot undertook successfully; others are based on problems that the pilot did not remedy. All of the lessons presented are discussed in the form of actions to undertake to yield positive results. It follows then, that not taking the actions may culminate in less than satisfactory results.

Two overriding lessons emerge from the set of lessons learned during the pilot study; they specifically relate to the establishment of a measurement process across multiple sites, domains, and application types:

- Those responsible for the initiation and implementation of a measurement process across multiple sites must be aware that every action has a ripple effect that is both broad and deep across the organizational units. It is therefore important to understand and follow the guidance and lessons learned from previous experience that are cited as necessary for success (e.g., [Basili], [Fenton], [Grady], [McAndrews], [Rifkin], [Rozum92], [Rozum93]). The value and importance of the lessons is magnified in direct proportion to the size and complexity of the organization.

- Organizational, communication, and personnel issues transcend the entire process and are just as significant, if not more so, than the technical issues related to software measurement. Paying attention to these issues will not guarantee success, but ignoring them will certainly result in a failed or less than satisfactory measurement process. This may be the most important lesson of all.

4.1 Organizational Goals and Motivation

For a measurement process to be effective and useful, it is essential that organizational objectives and goals be clearly stated at every level of the involved groups or units. Additionally, management support and participation, by review and oversight of the data, at

all levels is critical to achieving satisfactory results. Clear objectives and goals go hand-in-hand with management leadership and participation to provide the motivation, support, and energy required to initiate and sustain a software measurement process. The organizational objectives and management leadership provide the *raison d'être* for measurement, and without them, the measurements have little value.

In an operational sense, this means that the management at the highest level of the organization issues policies, directives, or instructions that express organizational objectives in terms of the products, processes, and resources, and identifies the attributes that are of interest or concern. The subordinate units, in turn, relate and elaborate on the objectives in terms of their mission and function, and initiate appropriate action (see below) across the organizational entities within their jurisdiction. Management continues to participate in the process by establishing oversight reviews at periodic intervals, giving energy and life to the policy visibly by its own actions. Merely issuing broad directives stating that the management supports software measurement or measurement is a policy that all must adhere to rarely makes any real or sustained progress towards obtaining measurement data that can be used to establish facts, make decisions, and take action.

4.2 The Measurement Team

Assigning the responsibility for implementing a measurement process is significant in that the reputation and influence of the persons selected send a clear signal to the entire organization relative to the seriousness of the effort. Ideally, an organization would select people who have a clear understanding of the purpose of the measurements, are trained in software measurement practices, and have experience in the organization which gives them credibility with the site management and staff. Realistically, this type of person either does not exist or is not available to work on the measurement team. In these cases, the person needs to be created by the organizations, and management support of them and the activity becomes even more critical. The sum of these effects is not only a positive signal to the organization, but more importantly, it can be a key factor in the success or failure of a measurement program.

An approach to implementing a measurement program similar to the one used by the pilot is to establish a group of motivated and trained personnel having the responsibility for all measurement activities. This approach has been proven successful many times in the past [Rifkin], [Rozum93]. In multisite situations, this often takes the form of a measurement process action team or working group that has the responsibility and authority to

- Establish measurement standards.
- Make decisions on which data need to be collected to obtain measurements that address the organization's and the projects' objectives.

- Consult with managers and personnel that must provide the data.
- Design the procedures for collecting, storing, and analyzing the data.
- Provide feedback to the site, and represent the site at the collaborative group deliberations.

Such a team might consist of senior managers and technical leaders who are given the authority and responsibility to introduce and enforce certain procedures or software processes necessary to make the required measurements.

4.3 Measurement Process Plans

It is important to devote resources and time at the outset of the measurement activity to create a plan for implementing and sustaining the measurement process. This is particularly important when establishing a measurement process for a large, multiple site or multiple project organization. Within the plan would be a schedule including time for identifying the scope of the effort clearly, as well as time to define and think through other aspects of the measurement process. In the pilot, a number of these activities were done in reaction to problems and issues as they occurred, as opposed to being done in a proactive manner up front.

The plan serves to communicate the processes for establishing the measurement objectives, defining the data to be collected, and developing procedures to be used to collect, analyze, and report the measurements. The plan also serves to describe the tasks that must be completed in order to implement the measurement process. Figure 4-1 illustrates a measurement process framework that might be used for creating such a plan [McAndrews].

Operationally, the plan for a large organization or multisite organization may well consist of several segments. The sponsoring organization may create that part of the plan that identifies the scope, and defines the data to be measured and the procedures to report the measurements. Other parts of the organization (e.g., a site or project that is responsible for collecting or storing data) should create the plans that define the procedures for collecting, verifying, and storing the data. A large organization with sites or units using different software development or maintenance processes will find this approach manageable, since the specific data collection procedures are likely to vary from site to site as a function of the differences in the software process.

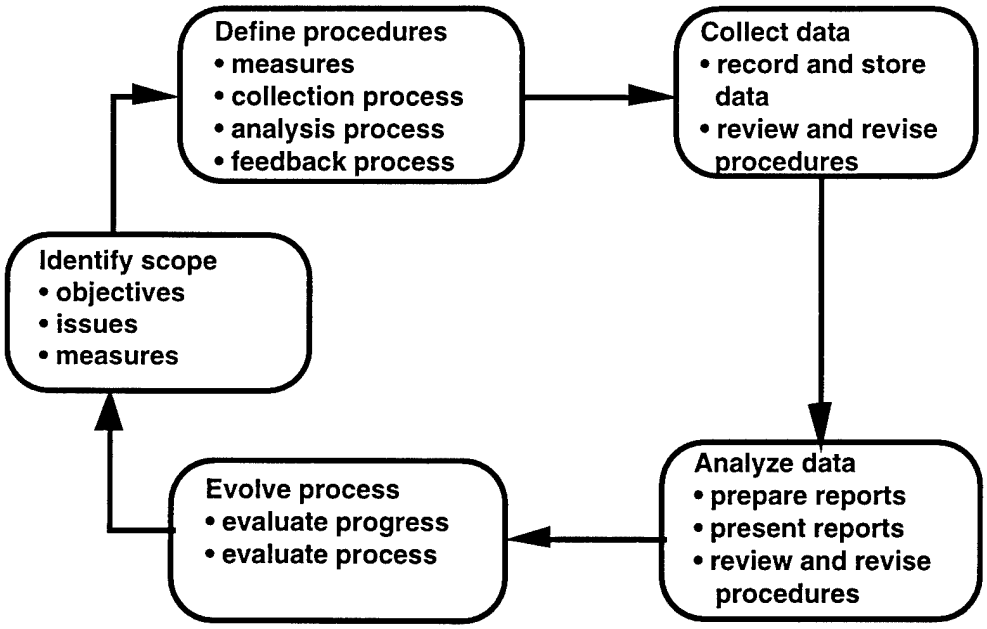


Figure 4-1: Framework for a Measurement Process

Other segments of the plan could consist of objectives, measurement definitions, procedures, etc., that are specific to the site or subordinate units. This provides the site an opportunity to integrate its measurement process with that of the sponsoring organization and ensure that objectives and measurements are consistent and supportive of the sponsoring organization.

4.4 Management Participation and Training

Management participation and support at all levels is essential for a successful measurement process. Those sponsoring and instituting a measurement process must consider and address the organizational and personnel issues that are frequently encountered in setting up a software measurement activity. Project management and staff within the organization will raise issues to upper management about the extra workload required, the measurements' relevancy to the software activities, the tendency to use the measurements to assess individuals, the use of the measurements by people who are not familiar with the software process, etc. Unless these issues are addressed satisfactorily at the outset, somewhere within the organization resistance to the process will arise or the entire measurement process will be discounted. The measurement process will then, in all likelihood, fail at some point in time.

There are three important lessons that have helped other organizations to address the issues outlined above (we did not, however, have the opportunity to observe these lessons during the pilot):

- Issue policy directives outlining the purpose and scope of the software measurement activity. These are useful for providing authoritative support and formally identifying the issues of concern to the organization.
- Ensure that management at all levels participate in the measurement process plan, conduct oversight reviews, and have a role in activities to set goals and objectives. This is necessary to gain their support and commitment, i.e., their buy-in. It is vital that they participate and see how they can affect the outcome, and how they can contribute to improving product, process, and resource measures.
- Provide training to all levels in the use of software measurement that shows how the measurements can and will be used to make decisions that improve product and process performance and meet the organization's objectives.

4.5 Piloting the Measurement Process

Conducting a pilot program to assess the effectiveness of a defined measurement process provides an opportunity to implement the process on a small or limited scale prior to broad implementation. Piloting the measurement process is intended to "test" the process without subjecting the entire organization to potential oversights or misunderstandings. The pilot should implement a defined measurement process according to a measurement plan on a small number of projects within the organization. Pilots based on informal and loosely defined procedures and processes, while perhaps feasible in a small, tightly knit organization, are prone to be futile due to the inability to extrapolate the procedures and processes to a much larger, multiple-site, multiple-domain organization. After a pilot is executed and the process refined based on the results, broad implementation of the measurement process can be planned and executed; however, during the early stages of the broad implementation, the measurement process should be expected to be further refined and will evolve based on a wider range of projects using the process.

4.6 Data Collection and Verification

Data collection and verification requires close attention to details and well defined procedures. If data collection is only for a single project or system, the effects of process misunderstandings or omissions may be limited to the project or system. This is not the case in a multiple-site, multiple-project measurement activity. The same misunderstandings or omissions ripple across all projects and take much longer to detect, and then correct and overcome. The measurement repository can be adversely affected if the data collection

requirements must be changed, further complicating the recovery. The following items were found to be useful to resolve these difficulties:

- Create unambiguous definitions of the data to be collected. Use the SEI checklists and supplementary forms to define data requirements, and use them to define data already being collected to provide a common reference point for existing data against required data. Be sure the data requirements are sufficient to provide the measurement users with the information they need to make decisions. Map the measurements to the data requirements to ensure that all the data required are defined in terms of how they are counted and when they are collected.
- Use the checklists as the basis for creating organization-specific data reporting forms. Ensure that the data report forms have entries corresponding to the checklists. Pay particular attention to the design, layout, and wording of the data report forms.
- Prepare a data collection guide that instructs the data collectors how to use the data report forms. Provide examples of filled out forms and include the definition checklists in the guide as background that data collectors can use to help answer their own questions. Include a glossary of terms and definitions, specifically, interpretations and translations from the checklists to the organization's software process.
- Exploit existing mechanisms (configuration control, project management, labor accounting, and problem tracking systems) to obtain the required data. Be prepared to modify or add software processes and administrative processes to obtain the required data.
- Provide and plan for enough time to install and revise data collection processes. Two delays will typically occur. The first delay typically occurs while organizations define how they will collect the data relative to their own software process and document how they have tailored the required definitions. Another delay typically occurs because not all organizations will have processes in place to collect the required data; some organizations will need to make changes to current processes or implement new processes to collect the data as defined.
- Establish verification, validation, and audit control procedures for the reported data. Initially this activity will be time consuming, but as the data collection process matures and trust and confidence is established in the data reported, the amount of time will decrease for performing this activity.

4.7 Common Measurement Definitions

The creation of common definitions of software measures sometimes, either explicitly or implicitly, assumes a common software process. This is particularly true if the definition is related to an event or presumed sequence of activities. In large organizations with multiple

projects and/or sites, processes are likely to be different among the various sites and projects. In such cases, if the data collection agent satisfies the data collection requirement by interpreting the definition to be consistent with the existing process (rather than changing the process), the data collected may have a different meaning than intended. While the SEI definition checklists and supplementary forms help identify many of these issues, they do not solve all of the problems that arise in this area.

If common definitions are not created, the differences in the data will be disguised; and analysis using the data could be misleading, depending on the analysis and degree of difference in the definition. During analysis, the importance of having common definitions and data that complies with the definitions affects comparing or aggregating the data when there is a mixture of projects or sites. In such circumstances, analysis of measures that includes data presumed to be commonly defined can lead to misinterpretation and incorrect conclusions.

Operationally, if the several organizational units are to report data in accordance with the common definitions, they may need to make changes to their software processes. If changing the processes to comply with the measurement definition is not feasible, this issue can be dealt with by providing means to characterize the measurement data. Examples would be creating and using data normalization techniques, collecting adequate information to characterize the project and process, establishing alternative measures that take the process differences into account, or providing the means for communicating the differences in the data during analysis.

4.8 Additional Lessons Learned

At the completion of the pilot software measurement effort, there were some sites and projects whose measurement capability was more mature than others. Since all the sites had the same training, used the same definition checklists, and received the same level of support and assistance, we have attempted to identify several of the key factors that appear to differentiate the sites with a more mature measurement process from those that struggled to implement a software measurement process consistently.

- *Site champions:* Site champions that fully appreciated software project management issues and thoroughly understood a project's software process are more apt to have success sooner and with fewer problems. Such site champions are frequently more proactive and connected to the site manager in a way that helped to influence the software measurement initiative in a positive manner.
- *Site and project initiative:* The sites that linked the measurement process and data collected to their business and/or mission-related goals and objectives appeared

more successful. Also, those projects that had their own goals and used software measures to determine status and progress were more successful than others.

- *Site management motivation:* Those sites whose functional and site executives provided project oversight using the measurement data to determine status and progress regularly, were more advanced in their software measurement capability than others.
- *Software process and supporting tools:* Those sites having some discipline within their software process (not necessarily documented) and commonalty of that process across projects, along with basic tools to support the software process (e.g., configuration management, problem tracking, labor tracking, schedule tracking), were able to make adjustments and establish a data collection process much more efficiently and effectively.

5. Conclusions and Recommendations

A large organization with multiple sites can write and implement a policy on software measurement that uses the SEI core measures as the basic units of measurement. Given what we have observed during this pilot, however, such a policy should be carefully crafted around specific purposes or objectives that the organization hopes to achieve with the measurement data collected.

For example, a policy may be implemented that is similar to the Hewlett Packard (HP) 10X program. In that program, HP issued a policy that its organizations would improve quality by a factor of 10. The effectiveness of any such policy issued will be dependent on the continual reinforcement of the policy objectives and regular review of the measures used to measure progress and performance toward meeting the policy objectives.

There are several methods of implementing such a policy. Two of the methods to implement a software measurement policy include

- Issue a policy or directive at the headquarters level (the Office of Secretary of Defense within the DoD) that establishes one or more general objectives and requires the various subordinate organizations to measure progress towards the objectives in a specific manner.
- Issue a policy or directive at the headquarters level that requires subordinate organizations with significant involvement in software acquisition, development, or maintenance to establish objectives relative to their own organizational goals, and to measure and report progress toward those goals. In this alternative, the affected subordinate organizations become the sponsoring organizations for specific policies that identify goals and objectives directly related to the organization and identify the means by which progress will be determined.

As evidenced by this pilot program and other efforts documented in the literature, successfully implementing such policies requires sufficient provisions for funding, staffing, training, and establishing a measurement process. It is recommended that the organization issuing the policy provide guidelines and standards that address these issues relative to the policy implementation, and identify supporting resources to provide guidance and assistance to the subordinate organizations.

However the policy is actually issued, if the policy includes the requirement to collect and report software measurement data, the organization issuing the policy needs to address the following:

- Define measurements to be adhered to by those reporting data. The definition can be articulated using a combination of SEI checklists describing what is to be

included and excluded in the data reported and textual descriptions to further describe aspects of data items that cannot be described by a checklist.

- Make use of the SEI checklists to communicate the results. The policy should use the checklists to define the data to be reported. Sites collecting the required data will need to describe how the data meet the policy's definition and possibly tailor or elaborate the policy definition with respect to the organization's software process.
- Develop, create, and provide a measurement process guide or handbook. The process guide should describe how and what the site can tailor to meet the requirements in the guide. The contents of a process guide should include
  - how the sponsoring organization will use the data to determine performance toward the objectives described in the policy;
  - the definition of the data contents to be collected (i.e., using the SEI checklists, describe what is to be included and excluded);
  - how the data are to be reported by a site to the organization, along with accompanying reporting forms for accomplishing this;
  - how a site can access the data stored at the organization level; and
  - how a site can expect to get feedback from the organization regarding its data.
- Identify the training and assistance available to the sites and how a site's point of contact can contact the organization issuing the policy. Recognize and provide the needed training for those developing the measurement process and also for those who will be expected to use the measurement results in their own software processes.
- Keep the reported data in a repository operated by the organization issuing the policy. Access may be unlimited, but the confidentiality of the data must be strictly enforced. That is, the data itself should be completely anonymous to those outside of the organizational chain. The database should include certain validity checks on the data prior to accepting the data into the repository.
- Require data to be reported electronically. The electronic transmission could be completed using a supplied template in the form of magnetic media or via telecommunication vehicles. This would require that the organization issuing the policy develop a reporting template and harmonize that template with the definitions and the database.
- Conduct pilot programs before attempting broad implementation.
- Recognize that not all sites will be able to comply with the policy immediately and a start-up period of time will be needed to implement the measurement process across the organization. The amount of time needed to comply will be a factor of the number of sites and projects involved. In a large, distributed organization, a year to bring all sites or projects into compliance with the new policy might not be unreasonable.

Bibliography

[Basili] Basili, V., and D. M. Weiss, "A Methodology for Collecting Valid Software Engineering Data," *IEEE Transactions on Software Engineering*, Vol. SE-10, no. 6 (Nov. 1984) pp. 728-738.

[Baumert] Baumert, John H. and Mark McWhinney, *Software Measures and the Capability Maturity Model* (CMU/SEI-92-TR-25, ADA 257 238), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, September 1992.

[Carleton] Carleton, Anita D., et. al., *Software Measurement for DoD Systems: Recommendations for Initial Core Measures* (CMU/SEI-92-TR-19, ADA 258 305), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, September 1992.

[Fenton] Fenton, Norman E., *Software Metrics: A Rigorous Approach*, New York, NY, Van Nostrand Reinhold, 1991.

[Florac] Florac, William A., *Software Quality Measurement: A Framework for Counting Problems and Defects* (CMU/SEI-92-TR-22, ADA 258 556), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, September 1992.

[Goethert] Goethert, Wolfhart B., et. al., *Software Effort & Schedule Measurement: A Framework for Counting Staff-Hours and Reporting Schedule Information* (CMU/SEI-92-TR-21, ADA 258 279), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, September 1992.

[Grady] Grady, Robert B., and Caswell, Debra L., *Software Metrics: Establishing a Company-wide Program*, Princeton, NJ, Prentice Hall, 1987.

[Humphrey] Humphrey, Watts S., *Managing the Software Process*, New York, NY, Addison-Wesley Publishing Co. 1989.

[McAndrews] McAndrews, Donald R., *Establishing a Software Measurement Process* (CMU/SEI-93-TR-16, ESC-TR-93-193), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, July 1993.

[Park] Park, Robert E., *Software Size Measurement: A Framework for Counting Source Statements* (CMU/SEI-92-TR-20, ADA 258 304), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, September 1992.

[Paulk] Paulk, Mark C., et. al., *Capability Maturity Model for Software, V1.1* (CMU/SEI-93-TR-24, ADA 263 403), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, February 1993.

[Rifkin] Rifkin, Stan, and Cox, Charles., *Measurement in Practice* (CMU/SEI-91-TR-16, ADA 241781), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, July 1991.

[Rozum92] Rozum, James A., *Software Measurement Concepts for Acquisition Program Managers* (CMU/SEI-92-TR-11, ADA 258 177), Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, June 1992.

[Rozum93] Rozum, James A., *The SEI and NAWC: Working Together to Establish a Software Measurement Program* (CMU/SEI-93-TR-07, ADA 280916). Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, December 1993.

Appendix A - Document Comment Forms for the SEI Measurement Checklists

Document comment forms were written for suggested changes to the checklists and accompanying forms for the SEI core measures. Those comments are contained in this appendix. The intent is to update the checklists and accompanying forms based on information from the DISA/CFSW pilot and others who have used the SEI core measures. Note that comments documented in this technical report and accompanying recommendations are not necessarily changes that will be made to the checklists; rather, they are suggested changes that have resulted solely from the DISA/CFSW pilot on the SEI core measures.

Comment No.: 1

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Develop a common header for checklists and report form example.

Checklists affected: All

Comment Description:

The checklists in the respective documents do not have header information that is consistent among them.

This problem was encountered by the DISA pilot metric projects, and led to confusion when using the checklists.

Recommended Solution and/or Rewording:

The DISA sites simplified the header information to include:

Site Name/ID

Project Name/ID

Date of Report

Reporting Period

Definition Name

This appeared to be adequate. Clearly, if an organization wanted to add more header type of information, it could do so, but it does not appear necessary to include more information than that above as a basic set.

Comment No.: 2

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Reorganize selections under the functional attribute of the effort checklist.

Checklists affected: Effort

Comment Description:

The *functional* attribute is arbitrarily divided into three categories: *product, build-level,* and *system level.* There is much duplication of attribute values among these three categories, and at the same time, some attributes are omitted from one or two categories. The DISA pilot sites found this confusing and unnecessarily confining. Some of the values listed under system or build were not listed under CSCI level, but the projects had no build or system-level activity. The functions are essentially independent of the level and are more dependent on the organizational situation.

Recommended Solution and/or Rewording:

It was suggested that the checklist be restructured so that all the functional activities are included under one attribute "functions" and another attribute be created to deal with the three (or more) levels, e.g., product, build, and system. Each of the levels, in fact, may contain all of the functions, but each product or project does not necessarily have multiple levels.

Comment No.: 3

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Add information to hour information attribute of effort checklist.

Checklists affected: Effort

Comment Description:

The *hour information* attribute does not provide for compensated time off during regular business hours. This is an often used category for employees who work extra hours (overtime, holidays, weekends, etc.) but are not paid extra; rather, they are given "comp" time from work to make up for the extra hours worked.

Recommended Solution and/or Rewording:

It was suggested that the checklist be restructured so that a third category be provided for salaried and hourly employees called administrative which would include comp time, vacation, illness, etc.

Comment No.: 4

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Modify employment class attribute information on the effort checklist.

Checklists affected: Effort

Comment Description:

The attribute *employment class* assumes all temporary employees and consultants are under contract. This frequently is not the case. There are temporary full time and temporary part time employees. Consultants generally work under contract, but also work part time and full time, similarly with subcontractors.

The language on page 15 of the effort and schedule measurement framework document and the checklist's categories needs to be clarified and amended regarding the counting of part-time and full-time staff-hours for these types of employees.

Recommended Solution and/or Rewording:

Revise the effort checklist to provide for four categories of employees: regular, temporary, consultant, and subcontract, each with the option of full time and part time. An alternative might be to establish a separate attribute for full time/part time, separate from employee class.

Comment No.: 5

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Add "original plan" column to the schedule checklist/report.

Checklists affected: Schedule

Comment Description:

The format of the "planned, changed, actual " data on the schedule reporting form and on the staff-hours reporting form was confusing. It would have been useful to include a column that contained the original plan date, in addition to the current format.

Recommended Solution and/or Rewording:

See comment description.

Comment No.: 6

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Modify effort and schedule checklists to include maintenance activities.

Checklists affected: Effort and schedule

Comment Description:

The effort and schedule checklists should be evaluated to determine their usability by a software project that is in maintenance mode. The issue is that scheduling is often based on a fixed schedule (e.g., whatever is ready by June 30 is shipped), or the individual change requests are scheduled, but the deliverables and milestones associated with the development cycle frequently are not employed and not required.

Recommended Solution and/or Rewording:

See comment description.

Comment No.: 7

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Add information to statement type attribute on the size checklist.

Checklists affected: Size

Comment Description:

There doesn't appear to be any straightforward way to identify LOC that are used to create screens, help files, and database schema other than the broad usage of "declarations."

Recommended Solution and/or Rewording:

Create additional subcategories to identify these segments of a program similar to the way comments are decomposed.

Comment No.: 8

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Move language-specific information to supplemental information of the size checklist.

Checklists affected: Size

Comment Description:

Beyond the first two pages of the checklist, the parts of the checklists and reporting forms that describe what was included or excluded should be included as supplemental forms. Typically, a project would only need to complete a small portion of these areas. Having them leads to confusion and questions of completeness.

Recommended Solution and/or Rewording:

See comment description.

Comment No.: 9

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Develop common column headers.

Checklists affected: All

Comment Description:

The terminology in the various checklists (data arrays versus specifications, include/ exclude, value count, etc.) is very confusing because the very different words and methods are meant to describe the same thing and be used the same way. The application of the checklists varies so much that it is difficult to use them in a comprehensive measurement program.

Recommended Solution and/or Rewording:

Standardize the column formats of the checklists. It is recommended that the format of the problem definition checklist be used, but that the column headers be more descriptive (i.e., array count, value count, etc.)

Comment No.: 10

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Replace reporting forms with discussion on tailoring/developing them.

Checklists affected: All

Comment Description:

The DISA/CFSW pilot projects were generally confused by the reporting forms. In general, because these types of forms are so severely tailored to the project, having a recommended form leads to confusion.

Recommended Solution and/or Rewording:

Delete the forms and discuss how one could be created and tailored to the project to match the checklist.

Comment No.: 11

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Reevaluate the shaded boxes that disallow their use.

Checklists affected: Quality and effort

Comment Description:

Reevaluate all boxes on the checklists that are shaded so that users don't check them. Often, those boxes would have been useful to the DISA/CFSW pilot. For example, the problem checklist had the testing boxes shaded all the way across under the *finding activity* attribute. The DISA pilot wanted to report separately, and by an array, all the problems/defects found during any kind of testing and those found in the field. The pilot had to be creative, when a simple modification to the form would have sufficed.

Recommended Solution and/or Rewording:

See comment description. Don't force a pattern to the completion of the checklists. Allow for hierarchical reporting when a project's process will support it and aggregate reporting for those that do not have the detailed software process definition.

Comment No.: 12

Originator's Name: DISA Pilot (SEI) Date Originated: August 94

Comment Title: Replace the DoD and DOD-STD-2167 terminology.

Checklists affected: All

Comment Description:

The checklists and supporting technical reports are very heavily ladened with DoD and DOD-STD-2167 terminology. However, only a (small) portion of the DoD actually uses this terminology.

Recommended Solution and/or Rewording:

Describe and use a more generic terminology. Don't be afraid to use a waterfall and structured programming terminology. Most people involved with software can relate to that terminology even if they do not use the method.

Appendix B - Checklists Used

The following pages include the actual checklists used by the DISA/CFSW measurement pilot to define the data collected and reported. For more information on the actual checklists and the details of how to use them, refer to [Goethert] for information on using the checklists for effort and schedule measures, [Park] for information on using the checklists for defining lines of code, and [Florac] for information on using the checklists for problem and defect measures.

Staff-Hour Definition Checklist

Definition Name: DISA Pilot - Total System Staff-hours

Date:

Originator:

Page: 1 of 3

	Include	Exclude	Report Totals
Type of Labor			
Direct	X		
Indirect		X	
Hour Information			
Regular time			X
Salaried	X		
Hourly	X		
Overtime			X
Salaried			
Compensated (paid)	X		
Comp Time	X		
Hourly			
Compensated (paid)	X		
Comp Time	X		
Employment Class			
Reporting organization			
Full time	X		X
Part time	X		
Contract			
Temporary employees	X		X
Subcontractor working on task with reporting organization	X		
Subcontractor working on sub-contracted task	X		
Consultants	X		
Labor Class			
Software management			
Level 1	X		
Level 2		X	
Level 3		X	
Higher		X	
Technical analysts & designers			
System engineer	X		
Software engineer/analyst	X		
Programmer	X		
Test personnel			
CSCI to CSCI integration	X		
IV&V	X		
Test & evaluation group (HW-SW)	X		
Software quality assurance	X		
Software configuration management	X		
Program librarian	X		
Data base administrator	X		
Documentation/publications	X		
Training personnel	X		
Support staff	X		

Definition Name: DISA Pilot - Total System Staff-hours

	Include	Exclude	Report Totals
Activity			
Development			
Primary development activity	X		
Development support activities			
Concept demo/prototypes	X		
Tools development, acquisition, installation & support		X	
Nondelivered software & test drivers	X		
Maintenance			
Repair	X		
Enhancements/major updates	X		
Product-Level Functions			
CSCI-Level Functions (Major Functional Element)			X
Software requirements analysis	X		
Design			
Preliminary design	X		
Detailed design	X		
Code & development testing			
Code & unit testing	X		
Function (CSC) integration and testing	X		
CSCI integration & testing	X		
IV&V	X		
Management	X		
Software quality assurance	X		
Configuration management	X		
Documentation	X		
Rework			
Software requirements	X		
Software implementation			
Re-design	X		
Re-coding	X		
Re-testing	X		
Documentation	X		
Build-Level Functions (Customer Release)			X
(Software Effort Only)	X		
CSCI to CSCI integration & checkout	X		
Hardware/software integration and test	X		
Management	X		
Software quality assurance	X		
Configuration management	X		
Documentation	X		
IV&V	X		

Definition Name:	Total System Staff-hours	Page: 3 of 3		
Product-Level Functions continued		Totals include	Totals exclude	Report totals
System-Level Functions				
(Software Effort Only)				
System requirements & design				
System requirements analysis			X	
System design			X	
Software requirements analysis			X	
Integration, test & evaluation				
System integration & testing			X	
Testing & evaluation			X	
Production and deployment			X	
Management			X	
Software quality assurance			X	
Configuration management			X	
Data			X	
Training				
Training of development employees			X	
Customer training			X	
Support			X	

Schedule Checklist
Part A: Date Information

Date: _____
Originator: _____

Project will record planned dates Yes _____ No _____
If "YES", reporting frequency Weekly _____ Monthly _____ Other _____
Project will record actual dates Yes _____ No _____
If "YES", reporting frequency Weekly _____ Monthly _____ Other _____

	Include	Exclude	Repeat each build	Relevant dates reported*
System requirements review		X		
System design review		X		
Software specification review	X			1
Preliminary design review	X			1
Critical design review	X		X	1
Code complete	X		X	1
Unit test complete	X		X	1
CSC integration and complete	X		X	5
Test readiness review/site readiness review	X		X	1
CSCI functional & physical configuration audits	X		X	1
Preliminary qualification test		X		
Formal qualification test		X		
Delivery & installation	X			10 *
Other system-level: Delivery to prime contractor		X		

Key to indicate "relevant dates reported" for reviews and audits

1-Internal review complete
2-Formal review with customer complete
3-Sign off by customer
4-All high-priority action items closed
5-All action items closed/deferred
6-Product of activity/phase placed under configuration management
7-Inspection of product signed off by QA
8-QA sign-off
9-Management sign-off
10-Shipped
11-_____

*Changed at 7 July meeting

‡ The DISA/CIM pilot only used 2 of the 3 pages.

Deliverable Products	Include	Exclude	Repeat each build	Relevant dates reported*
System-Level				
Preliminary system specification		X		
System/segment specification		X		
System/segment design document				
Preliminary Interface requirements spec.		X		
Interface requirements specification		X		
Preliminary interface design document		X		
Interface design document		X		
Software development plan		X		
Software test plan		X		
Software product specification(s)		X		
Software user's manual		X		
Software programmer's manual		X		
Firmware support manual		X		
Computer resources integrated support doc.		X		
Computer system operator's manual		X		
CSCI-Level				
Preliminary software requirements spec(s)		X		
Software requirements specification(s)	X			1, 6
Software preliminary design document(s)		X		
Software (detailed) design document(s)	X			2
Software test description(s) (cases)		X		
Software test description(s) (procedures)		X		
Software test report(s)	X			2
Source code	X			7, 1
Software development files		X		
Version description document(s)		X		

Key to indicate "relevant dates reported" for deliverable products

1-Product under configuration management
2-Internal delivery
3-Delivery to customer
4-Customer comments received
5-Changes incorporated
6-Sign-off by customer
7-Productivity
8-_____

Definition Checklist for Source Statement Counts

Definition name: DISA Pilot - Physical Source Lines of Code

Date:

Originator:

Measurement Unit:	Physical source line	X		
	Logical source statements			

Statement type	Definition	Data array		Includes	Excludes
When a line or statement contains more than one type, classify it as the type with the highest precedence.					
1 Executable	Order of precedence ->		1	X	
2 Nonexecutable					
3 Declarations			2	X	
4 Compiler directives			3	X	
5 Comments					
6 On their own lines			4		X
7 On lines with source code			5		X
8 Banners and nonblank spacers			6		X
9 Blank (empty) comments			7		X
10 Blank lines			8		X
11					

How produced	Definition	Data array	Includes	Excludes
1 Programmed			X	
2 Generated with source code generators			X	
3 Converted with automated translators			X	
4 Copied or reused without change			X	
5 Modified			X	
6 Removed				X
7				

Origin	Definition	Data array	Includes	Excludes
1 New work: no prior existence			X	
2 Prior work: taken or adapted from				
3 A previous version, build, or release			X	
4 Commercial, off-the-shelf software (COTS), other than libraries*				X
5 Government furnished software (GFS), other than reuse libraries*				X
6 Another product				X
7 A vendor-supplied language support library (unmodified)				X
8 A vendor-supplied operating system or utility (unmodified)				X
9 A local or modified language support library or operating system			X	
10 Other commercial library				X
11 A reuse library (software designed for reuse)			X	
12 Other software component or library			X	
13				
14				

Usage	Definition	Data array	Includes	Excludes
1 In or as part of the primary product			X	
2 External to or in support of the primary product				X
3				

* Anything that is modified or maintained would be included

Definition name:	DISA Pilot - Physical Source Lines of Code				
Delivery	Definition [X]	Data array [ ]		Includes	Excludes
1 Delivered					
2 Delivered as source				X	
3 Delivered in compiled or executable form, but not as source				X	
4 Not delivered					
5 Under configuration control					X
6 Not under configuration control					X
7					
Functionality	Definition [X]	Data array [ ]		Includes	Excludes
1 Operative				X	
2 Inoperative (dead, bypassed, unused, unreferenced, or unaccessed)					
3 Functional (intentional dead code, reactivated for special purposes)				X	
4 Nonfunctional (unintentionally present)				X	
Replications	Definition [X]	Data array [ ]		Includes	Excludes
1 Master source statements (originals)				X	
2 Physical replicates of master statements, stored in the master code				X	
3 Copies inserted, instantiated, or expanded when compiling or linking					X
4 Postproduction replicates as in distributed, redundant, or reparameterized systems					X
5					
Development status	Definition [ ]	Data array [X]		Includes	Excludes
Each statement has one and only one status,					
usually that of its parent unit.					
1 Estimated or planned					X
2 Designed					X
3 Coded					X
4 Unit tests completed*					X
5 Integrated into components					X
6 Test readiness review completed					X
7 Software (CSCI) tests completed					X
8 System tests completed				X	
9					
Language	Definition [ ]	Data array [X]		Includes	Excludes
List each source language on a separate line.					
1	Separate totals for each language			X	
2 Job control languages	_____			X	
3	_____			X	
4 Assembly languages	_____				X
5	_____			X	
6 Third generation languages	_____				
7	_____			X	
8 Fourth generation languages	_____				
9	_____				
10 Microcode	_____				X

*Changed at 7 July Meeting

Problem Count Definition Checklist-1

Software Product ID [ ]
Definition Identifier: [ ] Definition Date [ ]

Attributes/Values	Definition [ ] Specification [ ]			
Problem Status	**Include**	**Exclude**	**Value Count**	**Array Count**
Open	X			
Recognized				
Evaluated			X	
Resolved				
Closed	X		X	
Problem Type	**Include**	**Exclude**	**Value Count**	**Array Count**
Software defect				
Requirements defect	X			1
Design defect	X			1
Code defect	X			1
Operational document defect	X			1
Test case defect	X			
Other work product defect	X			
Other problems				
Hardware problem		X		
Operating system problem		X		
User mistake		X		
Operations mistake		X		
New requirement/enhancement		X		
Undetermined				
Not repeatable/Cause unknown		X		
Value not identified		X		
Uniqueness	**Include**	**Exclude**	**Value Count**	**Array Count**
Original	X			
Duplicate		X		
Value not identifed		X		
Criticality*	**Include**	**Exclude**	**Value Count**	**Array Count**
1st level (most critical)	X			1a
2nd level	X			1a
3rd level	X			1b
4th level	X			1b
5th level	X			1b
Value not identified		X		
Urgency	**Include**	**Exclude**	**Value Count**	**Array Count**
1st (most urgent)	X			
2nd	X			
3rd	X			
4th	X			
Value not identified		X		

*Level 1-5 represents MIL-STD-2167A definitions

Problem Count Definition Checklist-2

Software Product ID [ ]

Definition Identifier: [ ] Definition Date [ ]

Attributes/Values Definition [ ] Specification [ ]

Finding Activity	Include	Exclude	Value Count	Array Count
Synthesis of				
Design		X		
Code		X		
Test procedure		X		
User publications		X		
Inspections of				
Requirements		X		
Preliminary design		X		
Detailed design		X		
Code		X		
Operational documentation		X		
Test procedures		X		
Formal reviews of				
Plans		X		
Requirements		X		
Preliminary design		X		
Critical design		X		
Test readiness		X		
Formal qualification		X		
Testing				1a,1b
Planning	X			
Module (CSU)	X			
Component (CSC)	X			
Configuration item (CSCI)	X			
Integrate and test	X			
Independent verif. and valid.	X			
System	X			
Test and evaluate	X			
Acceptance	X			
Customer support				1a, 1b
Production/deployment		X		
Installation		X		
Operation		X		
Undetermined				
Value not identified		X		

Finding Mode	Include	Exclude	Value Count	Array Count
Static (non-operational)	X			
Dynamic (operational)	X			
Value not identified		X		

Appendix C - Reporting Forms Used

The following pages illustrate the actual data reporting forms used by the DISA/CFSW measurement pilot. There is no effort made in this report to describe how they were used by the pilot, nor is there a reference that one can use to obtain that information. The reporting forms are enclosed for information only and to illustrate how the reporting forms were tailored relative to the SEI examples included in the appropriate reference documents on each measure.

Version 2.1 March 24 1994
rsenamewestnrsa.wk1

DISA PILOT METRIC STAFF-HOURS REPORT FORM
Direct Staff-Hours Report

Site: _____ Date of Report: _____ (dd-mmm-yy)
Project: _____ Reporting Period: _____ (mmm-yy)
Release: _____

	Reporting Org		Contract		Total		Grand
	New, enh	Rework	New, enh	Rework	New, enh	Rework	Total
Development							
Regular time							
Overtime							
Comp			XXXXXX	XXXXXX			
Sub-total							
Maintenance							
Regular time							
Overtime							
Comp			XXXXXX	XXXXXX			
Sub-total							
Total Staff-hours							

OPTIONAL: HOURS

Requirements analysis			
Design			
Preliminary Design			
Detailed Design			
Code & Development Testing			
Code & Unit Testing			
Function Int. & Testing			
Integration & testing			
IV&V			
Management			
Software Quality Assurance			
Configuration Management			
Documentation			
Rework			
Software Requirements			
Software Implementation			
Re-design			
Re-coding			
Re-testing			
Documentation			
Other			

Version 2.1 Feb 2 1994
Filename=scnimm2.wk1

DISA Pilot Metric Schedule Reporting Form
Date Information
CSCI-Level Information

Site: _____ Date of Report: _____ (dd-mmm-yy)

Project: _____ Reporting Period: _____ (mmm-yy)

Release: _____

Milestones, Reviews & Audits*	Planned	Changed (yes/no)*	Actual
Software specification review			
Internal review complete			
Preliminary design review			
Internal review complete			
Critical design review			
Internal review complete			
Code Complete			
Internal review complete			
Unit test complete			
Internal review complete			
CSC Integration and test complete			
All action items complete			
Test readiness/Site readiness complete			
Internal review complete			
CSCI functional & physical config. audit			
Internal review complete			
Delivery & Installation			
Shipped			

* Only those completion criteria on the checklist appear below each deliverable. Enter yes or no in the "Changed" column if planned date has changed since last reporting period.

Version 2.1 Mar 7 1994
Filename=schrm.wk1

DISA Pilot Metric Schedule Reporting Form
Date Information
CSCI-Level Information

Site: _____ Date of Report: _____ (dd-mmm-yy)

Project: _____ Reporting Period: _____ (mmm-yy)

Release: _____

Deliverable Products*	Planned	Changed (yes/no)*	Actual
Software requirements specification			
Product under configuration control			
Sign-off by customer			
Software Design Document			
Internal delivery			
Software Test Report			
Internal delivery			
Software Code			
Product under configuration control			
Product delivery			

* Only those completion criteria on the checklist appear below each deliverable. Enter a yes/no in the "Changed" column if planned date has changed since last reporting period.

Version 2.1 14 March 94
Filename=prog.wk1

DISA PILOT METRIC PROGRESS REPORTING FORM

Site _____ Date of Report _____(dd-mmm-yy)

Project _____ Reporting Period _____(mmm-yy)

Release _____

Progress Data

	Total Planned	Changed (yes/no)	Actual To Date
# Units Designed			
# Units Integrated			
# Successful Test Cases			

Version 2.1 March14, 1994
Filename=probdfta.wk1

DISA PILOT METRICS PROBLEM/DEFECT REPORTING FORM

Site: _____ Date of Report: _____(dd-mmm-yy)

Project : _____ Reporting Period: _____(mmm-yy)

Release: _____

Opened During Current Reporting Period:

	Testing		Customer		
Type/Finding Activity	Crit 1-2	Crit 3-5	Crit 1-2	Crit 3-5	Total
Requirements defects					
Design defects					
Coding defects					
Documentation defects					
Other					
Total opened, this reporting period					

Closed During Current Reporting Period:

Type/Finding Activity	Crit 1-2	Crit 3-5	Crit 1-2	Crit 3-5	Total
Requirements defects					
Design defects					
Coding defects					
Documentation defects					
Other					
Total closed, this reporting period					