Software Engineering Institute

Measurement for Improvement: Successful Measurement Practices Used in Army Software Acquisition

James Wessel
Robert Ferguson

June 2009

TECHNICAL NOTE
CMU/SEI-2009-TN-008

Acquisition Support Program

http://www.sei.cmu.edu

Carnegie Mellon

This report was prepared for the

SEI Administrative Agent
ESC/XPK
5 Eglin Street
Hanscom AFB, MA 01731-2100

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.

This work is sponsored by the U.S. Department of Defense. The Software Engineering Institute is a federally funded research and development center sponsored by the U.S. Department of Defense.

This work was created in the performance of Federal Government Contract Number FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. The Government of the United States has a royalty-free government-purpose license to use, duplicate, or disclose the work, in whole or in part and in any manner, and to have or permit others to do so, for government purposes pursuant to the copyright license under the clause at 252.227-7013.

For information about purchasing paper copies of SEI reports, please visit the publications section of our website (http://www.sei.cmu.edu/publications/).

Table of Contents

Acknowledgments **v**

Abstract **vi**

1 Introduction **1**
- 1.1 Background of ASSIP 1
- 1.2 ASSIP Measurement Initiative 2
- 1.3 Scope of Effort 2

2 General Findings **3**

3 Examples of Successful Army Measurement Practices in Software Development **5**
- 3.1 Root Cause Analysis 5
- 3.2 Communication 5
- 3.3 Decision Analysis Measurement 5
- 3.4 Integrated Measurement Team 6
- 3.5 PMO Management 6
- 3.6 Quality Assurance 7
- 3.7 Requirements and Change Management 7
- 3.8 Risk Management 7

4 Enabling Successful Software Measurement Practices **9**
- 4.1 Foundational Elements 9
- 4.2 Academic Research 10
- 4.3 Army, DoD, and Industry Nexus 11

5 Conclusion **12**

Appendix A: Study Methodology **14**

Appendix B: ASSIP Measurement Survey **17**

Appendix C: Measure Interview Guide **18**

Appendix D: Army Acquisition Directives and Measurement **19**

Appendix E: Acronyms **20**

References **22**

List of Tables

Table 1: Program Survey Results 15

Table 2: Typical Measurement Data for Reporting 16

Table 3: Acronyms in This Technical Note 20

Acknowledgments

The authors gratefully acknowledge the representatives of the programs for participating in interviews, and completing and returning the survey. The authors also thank Cecilia Albert and Charlene Gross for their review of early drafts of this report. Their comments greatly improved the final product. Cecilia and Charlene are with the Carnegie Mellon® Software Engineering Institute (SEI).

Abstract

This technical note identifies and describes successful practices in software measurement that were discovered within a subset of current Army programs.

Conducted by the Carnegie Mellon® Software Engineering Institute (SEI) on behalf of the Army Strategic Software Improvement Plan (ASSIP), the study highlights software measurement practices that offices found to be valuable for problem identification, tracking, and active control of the program.

The intended audience for this report includes Army program managers, senior Army staff, program executive offices, software engineering centers, software engineering directorates, the Army Training and Doctrine Command (TRADOC), and the Army test community.

1 Introduction

Software is a primary enabler of today's military capabilities, and success in tactical operations is highly dependent upon reliable, effective, efficient, interoperable software. This technical note focuses on successful software measurement practices that Army acquisition organizations find to be valuable for software issue identification, tracking, and active control of programs. Conducted on behalf of the Army Strategic Software Improvement Plan (ASSIP), the study gathered information through a series of surveys and interviews within Army acquisition program management offices and software engineering centers.

Measurement is long chronicled as an enabler for acquisition improvement. For example, in the 2006 Defense Software Summit Report [DoD 2006a], the Program Executive Office (PEO) panel and plenary speakers summarized their perspective on software issues by stating that "improved systems and software engineering methods may reduce problem root causes and provide $24B in cost avoidance over the DoD Five-Year Defense Plan." In the same report, the Multi-Service and Defense Agency Panel recommended that strategic initiatives for software acquisition process improvement be established, and included program measurement as one of these strategic initiatives.

Measurement illuminates areas where an organization can focus efforts to significantly improve the speed of delivery and cost of software systems. It is for this reason that the Software Engineering Institute (SEI) was tasked by the Army to collaborate on a study of measurement activity and effectiveness throughout acquisition organizations and collaborators.

1.1 Background of ASSIP

ASSIP was established in 2002 by Army Acquisition Executive (AAE) Claude M. Bolton Jr. as a long-term partnership between the Army and the SEI. It was chartered to increase understanding of software acquisition program challenges, capture successful practices, and orchestrate initiatives toward the improved delivery of quality tactical software systems, on schedule and within budget.

An ASSIP Action Group (AAG) was chartered to help plan, coordinate, manage, and execute software acquisition improvement initiatives. The AAG membership currently includes representatives from the Army Program Executive Officers (PEOs), Software Engineering Centers and Directorates (SECs[1]), the Army Training and Doctrine Command (TRADOC), the Army Test Evaluation Command (ATEC), and the SEI.

[1] This document uses SEC to refer to any government software engineering support organization.

1.2 ASSIP Measurement Initiative

ASSIP places strong emphasis on its measurement initiative. The objectives for the ASSIP measurement initiative are to

- identify exceptional measurement practices within a subset of current army programs
- characterize the current state of program measurement within the Army
- quantify causal factors of SIS program issues that underlie chronic acquisition problems
- make recommendations to the Army based on successful practices

1.3 Scope of Effort

This report focuses on successful measurement practices that have been shown by either survey or observation to provide tangible value to the programs using them.

The team surveyed and interviewed Army program management offices (PMOs) to find successful measurement practices—ones that contributed to or enabled successful delivery of Army tactical capabilities. The team also reviewed the results of independent technical assessments of Department of Defense (DoD) programs by the SEI, GAO and other agencies. The scope of this activity spanned the involvement of PEOs and SECs, including

- PEO AMMO Ammunition (AMMO)
- PEO Aviation (AVN)
- PEO Command Control Communications Tactical (C3T)
- PEO Combat Support and Combat Service Support (CS&CSS)
- PEO Enterprise Information Systems (EIS)
- PEO Intelligence Electronic Warfare and Sensors (IEW&S)
- PEO Ground Combat Systems (GCS)
- PEO Soldier
- PEO Simulation, Training and Instrumentation (STRI)
- Fort Monmouth Communications-Electronics Life-cycle Management Command Software Engineering Center (SEC)
- Armament Software Engineering Center, Picatinny Arsenal SED
- Aviation and Missiles Research Development and Engineering Command (AMRDEC) SED, Redstone Arsenal, Alabama,
- Tank-automotive and Armaments Command (TACOM) SED

2 General Findings

*"Measurement **does not by itself improve the process;** however, the visibility it gives **provides insight** into planning, controlling, managing, and improving* [SEI 2006]*."*

The study team examined acquisition offices' management practices and approaches to software measurement that had proved to be effective. Effectiveness of measurement practices was determined by the delivery of the intended quality capability, on time and on budget, or the remediation of a program issue that posed a risk to on-time, on-budget delivery of the intended, quality capability.

The measurements practices described below were observed in more than one program having effective measurement practices, suggesting that they could be applicable in more than one context. Details of specific examples of measurement use are described in Section 3.

Measurement to plan and baseline

The team observed that core software measurement[2] data strengthens a program manager's ability to manage a software acquisition program. The study team was able to associate a program manager's ability to field a product successfully with his use of core measures to manage a program. Conversely, the team observed that the lack of measurement data and its use to manage the program office can contribute significantly to problems in program performance. This data deficiency also affects program managers' ability to plan, and contributes to their inability to assess the plans and performance of suppliers.

Measurements to evaluate plan versus actual

Some programs rely heavily on measurement to facilitate management within their PMO. These programs reported being able to avoid many systemic problems because of consistent and repeatable use of measurement as part of program management.

Measurements to estimate and predict

Programs with more comprehensive measurement data also had more robust plans and a clear understanding of the program's own performance. They had clear definitions regarding specific PMO roles, and each lead reported their conduct, supported by measurement data, within regular PMO meetings. Also, programs using measurement described recurring problems rather than reactively pinpointing problems one by one. This objective data supported their discussions and provided them with additional influence with both program manager (PM) and contractor.

Measurements to support process improvement

Another observed practice was measurement of the cumulative planned requirements changes. Planning for changes forces the program to allocate resources and to create processes to support the changes.

[2] In 1992, the SEI published a set of core software measures along with detailed checklists for specifying the data for these measures [Carlton 1992]. The four core measures identified were size, schedule, effort, and defects. The definition, meaning, and use of these measures have been subjects of continuous debate, but current research and investigation of both industry and academic literature provide strong support for their continued use.

3 Examples of Successful Army Measurement Practices in Software Development

This section describes actual events where software measurement data was successfully applied and illustrates the contribution of the measurement activity to that program. These examples can ultimately provide the foundation for broader implementation of measurement activities in the Army. The examples are categorized based on the type of effort and a description of the application of measurement practices.

3.1 Root Cause Analysis

Root cause analysis identifies causes of defects and other problems and points toward appropriate corrective actions. Instances observed by the study team highlighted the use of measurement to support objective identification of root causes and analysis to support successful issue management.

In one instance involving interface requirements, a program manager knew that there were schedule issues, but there was not enough data available to pinpoint the root cause of the schedule problem. Upon instituting a measurement program that included collections of requirements measures, a problem in the requirements area was discovered. A corrective action strategy based on objective data was put in place and ultimately led to successful management of the problem.

In another instance, data from core measures predicted a significant schedule delay, but this data alone was not sufficient to determine the root cause and take corrective action. When it was analyzed in conjunction with developer staffing data, the comparison showed that the delay was related to an anticipated dip in expert programmers. Contractor management arranged for greater cross training of its newer staff members to curtail the negative impact upon the program.

3.2 Communication

Using consistent data requirements, a program covers the progress of eight program projects using a common dashboard that is supported by regular data collection. The reports containing the data are generally available prior to the meeting—so the program meetings are not about reading the reports. A typical project discussion includes

- a summary of status—both technical and financial—by the program manager
- a review to cover the understanding of new or changed risks, issues and action items
- a discussion of plans for the next few weeks
- issues and action item assignments.

3.3 Decision Analysis Measurement

Decision analysis is a structured approach to evaluating alternative solutions against established criteria to determine a recommended solution to address an issue. Its formal structure is intended

to reduce the subjective nature of the decision and has a higher probability of selecting a solution that meets the multiple demands of relevant stakeholders [CMMI 2006].

One program office was observed to specifically require and use measurement to support decision analysis and resolution processes for proposal review and incentive payment. Measurement provided the objective evidence that could mitigate challenges to final decisions.

3.4 Integrated Measurement Team

The most effective uses of measurement observed occurred in programs where the program manager assigned a domain expert to validate, interpret, and be accountable for the respective domain area metrics. Thus, a finance person is responsible for interpreting finance measurement data, such as burn rate, funding, and expense. A system engineer is responsible for analysis of provided technical progress, risk, and product quality measures.

During the validation process, the domain expert partners with measurement experts to determine the most effective, efficient, and relevant measures to collect, analyze, and report. This is not a one-time process; periodic assessments are necessary to ensure relevance. Measurement indicators, the data collection, analysis, and reporting processes require adjustments within project phases and particularly when a program crosses a major milestone transition. Measures of progress are certainly different when transitioning from RFP work to design work. The nature of program risks also changes throughout the program life so the measurement reports need to reflect the risk profile of the project. Also, the program metrics and style of reporting are carefully crafted to suit the audience—enabling a full understanding of program status without overwhelming the audience with data.

For each of these measures or indicators, there is an expectation that trends and status can be understood as positive or negative. If the reported trend or status is negative, then the domain expert is expected to be prepared to explain “why.” The project team then determines whether the explanation suggested a risk or problem that needs to be addressed. Prioritized problems are formally tracked with an assigned owner.

Team members on programs that had better measurement data incorporated the behavior of individual accountability for specific measures or indicators. The individuals were queried about current data on a regular basis and were expected to describe how results compared to expectation. If there were differences, they were also expected to speak to potential impact and make the case for any required changes.

3.5 PMO Management

A program manager reported that collection for new measurement data suggested by a recent graduate of a software engineering course had reduced the number of issues and problems related to requirements management.

In another cases, an SEC provided a development team to the PMO where the development team was able to report using existing internal software performance measures. These measures allowed the PM to successfully guide the program through its acquisition life cycle.

3.6 Quality Assurance

Quality assurance need not be exhaustive. For efficiency, it is possible to sample in order to identify process compliance problems, provided there is some effort to identify the coverage and document the methods. The study team observed this sampling technique in one program. The development organization chooses a measurement and undergoes an extensive review of the information and personnel associated with the data.

3.7 Requirements and Change Management

Requirements management is an ongoing problem for many programs. Measurement of change can provide some indicators of the source of the problem. One employee was able contribute significantly to the development, understanding, and use of measurement for an Army effort after only one semester of a master's degree program. The first measurement indicators developed supported tracking requirements changes and the change management process. The program realized improved performance in this area, the program success was no longer endangered, and significant cost savings were reported as a consequence of the resulting improvements.

3.8 Risk Management

The majority of programs collect and use some measurement data in conjunction with risk management. Since "risk" is a comprehensive subject covering all aspects of a program, the level of detail and focus varies. The Army's "Probability of Success" program performance report was too new to be useful during this study, but it appears to provide some hope for program measurement data for future analysis.

Several programs were able to demonstrate their ability to identify risks and to mitigate the important ones. These programs had "burn down" charts showing that the risk management activities had reduced the overall risk profile of the program. This happens both by removing risks whose time has passed and by implementing an effective mitigation strategy.

Specific categories focused on software and software resources that the SEI team identified include the following:

- Key Performance Parameters (KPP)
  Risk mitigation includes architectural design reviews, third-party software assurance assessment, and extra software integration (evolutionary development).
- software resource availability
  Mitigation includes assignment of software engineer from an SEC, and additional training in software measurement.

At the software system development level, examples of observed key measurements used to illuminate latent software risks in programs are

- trouble reports (created, open, closed)
  risk: insufficient available effort to close defects before delivery

- source lines of code (delivered, planned)
  risk: cost of product or delivery will not match plan
- available and utilized resources
  risk: contractor lacks skilled resources to execute plan
- team level measures of schedule performance
  risk: a team falls behind or completes work too early
- code that fails to meet internal development standards
  risk: integration defects extend schedule
- measures of high complexity such as coupling and cyclomatic complexity
  risk: testing and maintenance rework will be costly
  risk: system will be fragile when exposed to changes implemented by other systems

4 Enabling Successful Software Measurement Practices

4.1 Foundational Elements

There are specific foundational elements that enable successful measurement practices. Previous discussion in this paper has shown that measurement practices have contributed to success in multiple software management areas. This section describes the foundational elements that were present and that, if more broadly implemented, could extend these successful practices across the Army.

As determined by software and acquisition communities over years of review and practice, the foundation elements for enabling effective measurement include

- evidence that leadership expects and will review measurement activities that are aligned with principal program objectives and support decision-making. These expectations are generally communicated via policy and regulations. Examples of DoD and Army measurement guidance are described in Appendix D, and illustrate that measurement can be an integral part of software management.
- an organizational structure which serves as a significant enabler of good measurement practice in successful programs. Organizational structures optimized for horizontal and vertical communication of information vital to program performance management are a key component. The primary mechanism to achieve this optimization was the IPT.
- a plan for performing measurement that assigns empowered responsibility and accountability, describes needed skill and training requirements, commits to measurement resources, and identifies communications with stakeholders. This foundational element supports consistent use and review of measurement, as seen in the program group that regularly reviewed the same items, risks, and issues in support of eight program projects. In both SEC organizations and selected PMOs, successful measurement programs required the presence of at least one person with specific training in measurement practice.
- a small number of trained staff (one or two) who assist with the definition of the measurement, the creation of the repository, and development of reporting procedures. The measurement staff has the responsibility to ensure the activities are implemented and the specific goals achieved. As such, these staff members help define the measurement indicators for communications, data collection, and reporting procedures. Measurement expertise helps programs be more effective developers and users of measurement data by participating in the following activities:
    a. a goal-driven measurement approach to facilitate the development of new indicators including developing and maintaining the documentation of the indicators and measurement practices
    b. careful and practical design of the data collection procedures to assure the integrity of the measurement system

c. periodic assessment of the quality and validity of the measurement data and the associated data collection procedures

d. reviewing new uses of existing indicators to assure the program that the representation meets the needs of the new measurement user

e. assisting the PMO with developing requests for additional supplier data and creating a useful representation of the data

- structure to manage the use of software measurement resources, such as data specifications, a measurement repository, analysis results, and tools enables collection and retention of valid historical data that is necessary for estimation, trend forecasts, and root case analysis. To this end the Army's Acquisition Information Management (AIM) database provides automated acquisition information tools to assist managers of Army acquisition programs to proactively manage assigned programs. The AIM database system is designed to allow the manager of each program to retain ownership of program data while providing access of this data to higher levels of the Army acquisition community. AIM provides a central repository to support statutory reporting requirements.
- approach to monitoring measurement efforts and applying quality assurance methods to enable program managers to ensure that measurement objectives align with organizational objectives. As seen in one of the examples, an established measurement approach provided schedule data but did not establish root cause for the problem. By incorporating staff data into the analysis, the cause of the problem was revealed and appropriate corrective action could be developed.

Virtually all the good news from the application of these foundational elements came from mature programs that have developed a working cadence in their product development. Programs that are still in the stages of proposal and planning have little to say about measurement even though there is a great deal of program office activity during these stages.

4.2 Academic Research

Most of the academic research is based on case studies, surveys, and detailed interviews with industry professionals. Of particular interest was the work of June M. Verner and William M. Evanco, who interviewed 21 senior project managers to devise a larger survey of 122 projects [Verner 2005]. The research shows the following:

- The greatest opportunity for improvement is at a project's start in the requirements and risk identification and control areas. User communities that provide adequate time and resources to requirements definition correlate well with program success.
- Project management success is highly dependent on the experience and skills of the project manager.
- Postmortem reviews are important for process improvement, but companies seldom perform them. As a result, the companies tend to repeat mistakes.
- Good cost and schedule estimates affect project success. The best estimates came from those with training and experience in estimation. The worst estimates came from senior manage-

ment, marketing and the customer. Seventy-one percent of estimates that were considered "good" involved the participation of an individual with project management experience.

- The release decision was made with sufficient information about validated requirements.

4.3 Army, DoD, and Industry Nexus

In general, the findings from industry and other DoD service measurement investigations are in agreement with the findings of this initiative. The important difference is the context, in this case the Army acquisition environment. The team's findings show that the same measurement principles successfully employed within industry and other DoD services can be effective in the Army acquisition context.

Successful PM offices that used measurement had several trends in common. These trends matched those in industry in many significant areas:

- They provided a fundamental core measurement framework to support good measurement activities.
- They educated management staff in practical measurement use.
- They held staff accountable for measurement-related action items.
- They organized their program to support measurement activity.
- Successful staffs invested in a measurement process where all stakeholders regularly met and managed program objectives by measurement data collected and analyzed.
- They used or developed tools and automated processes to aid measurement processes and accommodate the decision process.
- Many offices leveraged the measurement expertise resident in the local SED/SEC organizations. Some of these supporting staff members were embedded into the program environment.
- Acquisition teams consisting of acquisition office, government staff, developer organizations, SED/SEC organizations worked as a seamless organization.
- Staff members enjoyed working together and gained significant satisfaction carrying out their respective duties.
- Horizontal and vertical communication channels within the program were open. All staff members understood their roles and those of staff members in other channels of the program.

5 Conclusion

This initiative conducted a search for software system acquisition measurement practices within acquisition databases, websites, periodicals, news articles and other resources. AAG members reviewed their respective organization's current repositories and archives for measurement data, as did other Army organizations (such as Office of Deputy Assistant Secretary of the Army-Cost and Economics (ODASA-CE)).

Managers across multiple programs seem to connect the use of measurement with success in baselining and planning, evaluation of progress and prediction, and overall improvement. Individual examples of measurement and analysis indicate that measurement practices can support efficiency and successful fielding in a variety of contexts. Areas of particular strength that were observed were root cause analysis, integrated teaming for measurement and domain expertise, and risk management.

One or more foundational elements that enable successful measurement practices were present in all cases presented in this paper. Broader implementation of the foundational elements will allow the Army to extend successful measurement practices across the organization. These elements included

- evidence that leadership expects and will review measurement activities that are aligned with information needs and support decision-making
- a plan for performing measurement that assigns responsibility and accountability, describes needed skill and training requirements, and identifies communications with stakeholders
- managing the measurement data specifications, storage (such as a measurement repository), analysis results, and tools that enable collection and retention of valid historical data that is necessary for estimation, trend forecasts, and root case analysis
- monitoring of measurement efforts and applying quality assurance methods to enable program managers to ensure that measurement objectives align with organizational objectives

The study team found that successful PM offices that used measurement had several trends in common, which matched those in industry in many significant areas:

- They provided a fundamental core measurement framework to support good measurement activities.
- They educated management staff in practical measurement use.
- They held staff accountable for measurement-related action items.
- They organized their program to support measurement activity.
- Successful staffs invested in a measurement process where all stakeholders regularly met and managed program objectives by measurement data collected and analyzed.
- They used or developed tools and automated processes to aid measurement processes and accommodate the decision process.

- Many offices leveraged the measurement expertise resident in the local SED/SEC organizations. Some of these supporting staff members were embedded into the program environment.
- Acquisition teams consisting of acquisition office, government staff, developer organizations, and SED/SEC organizations worked as a seamless organization.
- Staff members enjoyed working together and gained significant satisfaction carrying out their respective duties.
- Horizontal and vertical communication channels within the program were open. All staff members understood their roles and those of staff members in other channels of the program.

Appendix A: Study Methodology

Develop Baseline of Current Activity

The first step in this initiative was to develop a strategy to collect and analyze program software measurement data currently in use across Army programs. The initial intent was to collect available core software measurement data from SIS programs toward developing a common baseline for trend analysis across programs, and within particular domains or portfolios. With more information needed, the team moved to a process of PMO interviews and on-site assessments to identify effective measurement practices in software intensive programs. The original team comprised the authors, James Wessel and Robert Ferguson of the SEI, and PEO and SEC representatives from the ASSIP.

The team obtained data on SIS acquisition program measurement and process information that was spread organizationally and geographically across PEO and program manager offices within the Army. The Acquisition Information Management (AIM) database was sparsely populated, and there did not appear to be any other central data source. The team's efforts then turned to gaining access to this information directly from PEO or program manager offices using interviews and surveys. An AAG Software Measurement integrated product team (IPT) was established to support these and other activities related to the study.

The study team performed a cursory review of software measurement practice within academia, industry and other DoD service organizations. To establish the breadth of Army measurement directives (as opposed to actual practices), the team also gathered a compendium of Army policy, regulation, guidelines, and memoranda regarding the conduct and use of software measurement. It is presented in Appendix D, Army Acquisition Directives and Measurement.

Collecting Artifacts

The study team collected and analyzed artifacts that the programs use in their measurement efforts, including tools, methods, contextual information, processes, documentation, validation, use, and interpretations in each instance. Of particular importance was the contextual information, because, in practice, measurement data standards vary across the Army's programs.[3] The contextual information is important in deciphering the material for program managers and other users of the data and also eases the adoption of the practice by groups with similar needs.

[3] The Army has, however, issued numerous directives regarding measurement (see Appendix D, Army Acquisition Directives and Measurement).

Analysis of Results

For this study, there were actually multiple data calls to programs and PEOs over a three-year period. In total there were survey responses from six PEOs and 15 programs.

Table 1: Program Survey Results

Number of programs responding	What data do you collect?
6	Software size (new, modified, reused)
1	Software complexity
4	Productivity in lines of code (LOC)/staff-month
3	Software Resources Data Report (SRDR)
5	Defect removed (or defect density)

Additionally, the software engineering centers (SEC/SED organizations) had measurement of certain key processes involving defect removal. These software engineering centers had a good relationship with the program office.

The productivity (LOC/effort) measure was used primarily to show the center had been able to improve its productivity over time.

The use of the SRDR was also interesting. The SRDR provides primarily project classification and historical measurement data that is a result of estimates and software development. It is used in executive reporting and could be used as a means of reporting trends across collections of programs. It is not designed to provide any sort of information that can be used in program control and risk management. The SRDR is intended to support Army program measurement in the longer term. The study team believes the SRDR can provide a valuable basis for program estimates. SRDR values collected are shown in Table 2.

Table 2: Typical Measurement Data for Reporting

Measurement data	Description
Project	Name, version, date, authorization (memorandum of understanding, contract, etc.) and reporting event
Organization	Development organization, CMMI level, evaluator, date of certification
Product Description	• Top 4 application types and percent of feature content • Development process, new/upgrade • Top 2 languages used • Commercial off the shelf/government off the shelf used • Peak staff • Staff (highly experienced, nominal experience, entry level)
Product Size	• Number of requirements • Number of external interfaces • Requirements volatility (1=low, 5=high) • New code size • Modified code size • Unmodified and reused code size
Resource/Schedule Reporting	Duration and effort for the following activities • software requirements • software architecture and design • software coding and unit test • software integration and system/software integration • software qualification testing • software developmental test and evaluation • all other direct software engineering development
Product/Quality Reporting	Mean time to serious or critical defect

In terms of program management, the information that has been selected for the SRDR is appropriate to several application categories that were identified in the specific indicators section above:

- quality assurance
- high-level risk assessment
- estimation
- allocation of budget resources

Appendix B: ASSIP Measurement Survey

Survey

The purpose of the measurement survey was the effective and efficient collection of program software measurement information. The survey was addressed to all AAG PEO representatives to obtain software measures, and associated contract language, reported by contractors to their respective Acquisition Category (ACAT) Level I (ID and IC) and II PMO offices.

The survey was designed to

- collect information on current utilization of measurement information within programs
- uncover some potential measurement needs
- identify examples of useful contract language

The initiative team used the following survey questions with the Program Managers.

Introductory Text

As you are aware, under ASSIP, we are seeking to identify useful practices related to software measurement in order to offer improvement and suggestions for contracting and oversight. SEI has interviewed a few PMs and found positive results. The next step is to try and establish a "baseline" of measurements practices across a wider sample of Army programs. We need your help.

Respectfully ask PEO Reps to report at the Sep AAG meeting results of the following survey for ACAT I and II programs:

1) What metrics for software are reported to the PMO by your contractor (e.g., size, effort, schedule, quality)? Other?
2) How often are software metrics reported - weekly, monthly, quarterly, other?
3) Please provide a copy of the most recent report.
4) Please identify a POC.

If we find that the practices in your program are helpful, we would like to further interview your POC to see exactly how your measurements have helped to facilitate contracting and oversight.

Appendix C: Measure Interview Guide

The initiative team used the following text to guide the phone interview with the program manager or a senior deputy.

Introduction

Program failure has sometimes been attributed to a lack of measurement data by senior leadership. Since this has been cited as a problem for over a decade and continues to be a problem today, it seems likely that we have not correctly identified the problem or its causes. The SEI is interested in gaining a better understanding of the problem and its causes.

1) Suppose the SEI had some measurement expertise and were to visit. What might you like to tell them about measurement problems and proposals within your program office?
2) Describe the measurement and metrics that you feel are most valuable for running your program. Are you satisfied that you get this data in a timely fashion?
3) Do these measures help you discover, understand and mitigate your top program risks during each phase of your acquisition life cycle?
4) Do you believe that the system used to get the measurement data is giving you a consistent story and a reliable picture of things that are important to you (relative to your product and your program)?
5) A major part of running a program is eliciting and making promises with other parts of the Army. Do you feel that you have useful data to help you negotiate these promises? Would having better measurement and better ways to tell your story help? Describe a sample problem or issue where you feel better measurement data would have helped.
6) A better understanding of the supplier's data and presentation of measurement might make the staff of the program office shrewder in identifying possible problems in the suppliers. Do you have a story about people misinterpreting the supplier's data?
7) Some measurement and analysis of the external forces might help the program manager negotiate better commitments from other stakeholders. For example, it is sometimes difficult to obtain GFE/GFI in a timely fashion. Sometimes the attendance to or response to official program reviews does not meet program promises. Has this type of issue been a problem? Can you cite a better example of the problem?
8) Program offices tend to be understaffed and very busy. It might happen that people are not working on the important things and that many important things take too long or require significant rework. Some monitoring of the activities of the program office might identify ways to improve processes and resource allocation. Would this be a concern for your PMO? Again if you could cite an example that concerns you, it would be valuable to help with the construction of additional questions and making a proposal to work for your program.

Appendix D: Army Acquisition Directives and Measurement

The study team reviewed applicable policy, regulation, instruction, and guidelines for Army acquisition. Listed below are examples of acquisition measurement mandates discovered during this study.

DOD GUIDANCE

The *Defense Acquisition Guidebook* (DAG), version 1.6, requires program managers to develop program metrics as part of its Systems Engineering Plan [DoD 2006b]. The DAG additionally states that it is the program manager's responsibility, as the program life-cycle manager, to develop program metrics, as stated in DAG Section 5.1.3.

Some very specific DoD measurement instruction currently exists for a Net-Ready Key Performance Parameter (NR KPP) in DoD Instruction 4630.8, *Procedures for Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS)* [DoD 2004].

DoD Instruction 5000.2 requires that developmental test and evaluation objectives be structured to provide accurate, timely, and essential information to decision makers throughout the system life cycle. It is further mandated that the program manager shall prepare a SEP for each milestone review to include a description of applicable metrics. [DoD 2008].

Army Regulation

Army Regulation (AR) 70-1, Army Acquisition Policy (dated December 31, 2003), Section 7-13, requires program managers to negotiate software metrics with the developer to bring about necessary discipline in the software development process and to assess the maturity of the software product. A minimum set of metrics is recommended as mandatory for programs to collect and use during the conduct of managing their programs [HQDA 2003a].

The May 30, 2003 version of the Army (DA) Pamphlet 73-1, Appendix Q, provides a summary of "software areas of interest and potential measures." In addition, Section VI—*Army Software Metrics*—"provides 14 examples of software metrics that can be used to gather information on the status of software throughout the life cycle of Army software-intensive systems [HQDA 2003b]."

Army Memoranda

An Army policy to specifically address metrics was issued on September 19, 1996, by the Director of C4 Information Systems titled *Acquisition Reform and Software Metrics*. The policy directs program managers to "take advantage of those metrics that are part of the developers' normal business practice." [DoArmy 1996] The memo reaffirms required metrics previously cited in DoD 5000.2-R, *Mandatory Procedures for Major Defense Acquisition Programs (MDAPs) and Major Automated Information System (MAIS) Acquisition Programs*, to include management-related program metrics (memorandum reference a, appendix V) [DoD 2002].

Appendix E: Acronyms

This report also contains many acronyms, which, for ease of reading, are defined in the following table.

Table 3: Acronyms in This Technical Note

Acronym	Description and Definition
AAE	Army Acquisition Executive (see also ASA(ALT)/AAE)
AAG	ASSIP Action Group; chartered as a multidisciplinary team with a charge to generate and provide guidance to software-intensive systems improvement initiatives
ACAT I	Acquisition Category I; a major defense acquisition program (MDAP) subject to Defense Acquisition Board oversight and estimated by the Under Secretary of Defense (Acquisition, Technology and Logistics) to require an eventual total expenditure of more than $365 million in research, development, test and evaluation funds, or $2.190 billion in procurement funds measured in FY2000 constant dollars.
ACAT II	Acquisition Category (ACAT) II; programs are acquisition programs that do not meet the criteria for an ACAT I program, but do meet the criteria for a major system. These programs are managed by a program manager who reports to a PEO or a materiel command as designated by the Army Acquisition Executive (AAE). These programs receive an Army Systems Acquisition Review Council (ASARC) review and require a decision by the AAE at each milestone review.
ACAT III	Acquisition Category (ACAT) III; programs are non-major programs (including non-major AIS programs) that are designated by the AAE or the Army Chief Information Officer (CIO), due to special interest and are managed by a program manager who reports to a PEO or a materiel command as designated by the AAE or CIO. These programs receive an in-process review (IPR) and require a decision by the PEO or the commander of the materiel developing command at each milestone review point.
AIM database	Acquisition Information Management database
Army Measurement IPT	ASSIP Measurement Integrated Product Team
ASA(ALT) MILDEP	Military Deputy to The Assistant Secretary of the Army for Acquisition, Logistics and Technology
ASA(ALT)/AAE	Assistant Secretary of the Army for Acquisition, Logistics and Technology/Army Acquisition Executive; the ASA(ALT) responsibilities include appointing, managing and evaluating program executive officers and program managers; managing the Army Acquisition Corps; and overseeing research, development, test, evaluation and acquisition programs. For more information, see https://webportal.saalt.army.mil/main/aae.htm

Acronym	Description and Definition
ASA(FM)	Assistant Secretary of the Army for Financial Management
ASSIP	U.S. Army Strategic Software Improvement Program; chartered to improve the acquisition of software-intensive systems
ATEC	Army Test Evaluation Command
DAG	Defense Acquisition Guidebook
DAU	Defense Acquisition University
DID	data item descriptions
DoD	U.S. Department of Defense
DRPM	Director Reporting Program Manager
GAO	U.S. General Accounting Office
OTEC	Operational Test and Evaluation Command
PEO	program executive officer
PM	Army program managers
PMO	Program Management Office
P(S) report	probability of success report
QA	quality assurance
SEC	Software Engineering Center
SED	Software Engineering Directorate
SIS	software-intensive systems
SLOC	source lines of code
SoS	system of systems
SPMN	Software Project Manager Network
SRDR	Software Resources Data Report
SSG	Senior Steering Group; a senior ASSIP governance body
SSIMP	Strategic Software Improvement Master Plan
STEP	Army Software Test and Evaluation Panel
TR	trouble report
TRADOC	U.S. Army Training and Doctrine Command

References

[Carlton 1992]
Carlton, Anita D.; Park, Robert E.; Goethert, Wolfhart B.; Florac, William A.; Bailey, Elizabeth K.; & Pfleeger, Shari Lawrence. *Software Measurement for DOD Systems: Recommendations for Initial Core Measures.* (CMU/SEI-92-TR-019, ADA258305). Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University, 1992.
http://www.sei.cmu.edu/publications/documents/92.reports/92.tr.019.html

[CMMI 2006]
CMMI Product Team. *CMMI for Development, Version 1.2*. (CMU/SEI-2006-TR-008, ADA453524). Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University, 2006.
http://www.sei.cmu.edu/pub/documents/06.reports/pdf/06tr008.pdf

[DoArmy 1996]
Department of the Army, Director of Information Systems for Command, Control, Communications, and Computers. *Memorandum: Acquisition Reform and Software Metrics*. September 19, 1996.

[DoD 2008]
Department of Defense. *Operation of the Defense Acquisition System* (DoD 5000.2). Washington, D.C.: Office of the Undersecretary of Defense (Acquisition, Technology, and Logistics), 2008.
http://www.dtic.mil/whs/directives/corres/pdf/500002p.pdf

[DoD 2006a]
Office of the Deputy Under Secretary of Defense for Acquisition and Technology and National Defense Industrial Association. *Defense Software Summit Report (Final Draft, Version 1.0).* October 2006.Washington, DC.

[DoD 2006b]
Defense Acquisition University. *Defense Acquisition Guidebook*. (2006)
http://akss.dau.mil/dag/DoD5000.asp?view=document

[DoD 2004]
Department of Defense. *Procedures for Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS).* (Dodo 4630.8). Washington, D.C.: Office of the Assistant Secretary of Defense for Networks and Information Integration, 2004.
http://jitc.fhu.disa.mil/jitc_dri/pdfs/i46308.pdf

[DoD 2002]
Department of Defense. *Mandatory Procedures for Major Defense Acquisition Programs (MDAPs) and Major Automated Information System (MAIS) Acquisition Programs.* (DoD 5000.2-R). Washington, D.C.: Office of the Undersecretary of Defense (Acquisition, Technology, and Logistics), 2002.
http://www.explorationsystems.nasa.gov/documents/TTT_052005/DoD50002R.pdf

[HQDA 2003a]
Headquarters, Department of the Army. *Army Regulation 70-1: Army Acquisition Policy*. (2003) http://www.army.mil/usapa/epubs/pdf/r70_1.pdf

[HQDA 2003b]
Headquarters, Department of the Army. *Department of the Army Pamphlet 73-1: Test and Evaluation in Support of Systems Acquisition.* (2003) http://www.usapa.army.mil/pdffiles/p73_1.pdf

[SEI 2006]
Software Engineering Institute. Course "Implementing Goal-driven Measurement." 2006.

[Verner 2005]
Verner, June M. & Evanco, William M. "What Project Management Practices Lead to Success?" *IEEE Software* 22, 1 (January/February 2005): 86-93. http://doi.ieeecomputersociety.org/10.1109/MS.2005.12

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
(Leave Blank)	June 2009	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Measurement for Improvement: Successful Measurement Practices Used in Army Software Acquisition	FA8721-05-C-0003

6. AUTHOR(S)

James Wessel

Robert Ferguson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	CMU/SEI-2009-TN-008

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
HQ ESC/XPK 5 Eglin Street Hanscom AFB, MA 01731-2116	

11. SUPPLEMENTARY NOTES

12A DISTRIBUTION/AVAILABILITY STATEMENT	12B DISTRIBUTION CODE
Unclassified/Unlimited, DTIC, NTIS	

13. ABSTRACT (MAXIMUM 200 WORDS)

This technical note identifies and describes successful practices in software measurement that were discovered within a subset of current Army programs.

Conducted by the Carnegie Mellon® Software Engineering Institute (SEI) on behalf of the Army Strategic Software Improvement Plan (ASSIP), the study highlights software measurement practices that offices found to be valuable for problem identification, tracking, and active control of the program.

The intended audience for this report includes Army program managers, senior Army staff, pro-gram executive offices, software engineering centers, software engineering directorates, the Army Training and Doctrine Command (TRADOC), and the Army test community.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Acquisition, software-intensive systems, metric, measurement	33

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL