Software Capability Evaluations: Experiences from the Field

Abstract

The concepts of software process maturity developed by the Software Engineering Institute (SEI) at Carnegie Mellon University are being used by a growing number of U.S. government and industrial software organizations. This paper gives an overview of the Software Capability Evaluation (SCE) method used by government agencies to select contractors and monitor contracts, using the SEI Process Maturity Framework as a yardstick. The paper also discusses the evolution of the SCE method since its introduction in late 1987.
ACKNOWLEDGEMENTS

This paper would not have been possible without the work of Ed Averill, Watts Humphrey, Albert Johnson, and Jim Withey. Our efforts in developing and refining software capability evaluations also rely on contributions from many people in government and industry. Their support is greatly appreciated. We also thank Jim Armitage, Pat Hanavan, and Marilyn Bush for providing stimulating comments. The authors are indebted to Suzanne Couturiaux for her thorough and professional editing of the paper. Finally, though he will accept no credit, we thank Bob Kent for making it all happen.
1 INTRODUCTION

This paper reviews the Software Engineering Institute (SEI) Software Capability Evaluation (SCE) method, explores the concepts and methods of SCE, and describes some experiences and lessons learned from applying and observing the use of SCE in Department of Defense (DoD) contracts, primarily those in the Air Force. This paper emphasizes the practical applications of SCE so that users can better understand how it applies in an acquisition process. It will help government agencies that are considering adding SCE to their acquisition process and industry software developers who are preparing for an SCE. The paper does not teach how to conduct an SCE, but explains what happens in a typical SCE.

The paper provides an overview of the following topics:

- the role of SCE in software process improvement
- when an SCE is appropriate in an acquisition process
- how to plan and execute an SCE
- how to use the findings from an SCE
- issues facing the SCE method
- evolution of the SCE method

We assume readers are familiar with the government acquisition process, as well as with the process maturity framework developed at the SEI. For more information on the framework, see CMU/SEI-87-TR-11 [Humphrey 87a].
THE ROLE OF SCE IN SOFTWARE PROCESS IMPROVEMENT

Software schedules, quality, and performance have repeatedly failed to meet the expectations of managers from both government and industry. This is especially true in the DoD, where weapons software systems are larger and more complex than most commercial applications. At the same time, the systems being acquired by the government have increasingly grown to rely on software.

From these problems, the government has recognized a need to gain insight into the software process capability of organizations bidding on contracts because no contractor can perform predictably and consistently at a level of software process maturity beyond its inherent capability.

As part of its mission, the SEI has focused on improving the software process. The software process includes those processes, methods, and tools that establish the managerial and technical environment in which software products are developed, for example, design reviews, project management, and quality assurance activities. The importance of the software process to a project’s technical and managerial success increases with the size and complexity of the software product being developed.

The SEI and the MITRE Corporation have jointly developed a maturity framework\(^1\) to measure the maturity of a developer’s software process [Humphrey 88], [Humphrey 89]. The principle behind the maturity framework is that the quality of a product stems, in large part, from the quality of the process used to create it. To consistently improve products, the process used to develop them should be understood, defined, measured, and progressively improved. Concepts of process and quality management [Crosby 79], [Deming 86], [Juran 88] are applied to building and maintaining software products through the maturity framework.

Five basic levels of process maturity are defined in the model to describe the progression from an ad hoc software process to one that is under statistical control and that can act as a stable foundation for continuous process improvement. The maturity levels are characterized by key process areas that focus on improvement efforts at each maturity level. All key process areas at one maturity level must be satisfied before an organization can effectively advance to the next maturity level. Examples of key process areas include Software Quality Assurance to achieve level 2, and Training to achieve level 3.

The maturity framework is applied in two SEI process evaluation methods:

1. The Software Capability Evaluation (SCE) method, an independent evaluation of the software process of an organization to gain insight into its software development capability. The SCE is funded by the government agency awarding the contract.
2. The Software Process Assessment (SPA) method, an assessment by an organization of its own software process to measure its process maturity level, to gain insight into its software development capability, and to prioritize actions for software process improvement.

\(^{1}\) CMU/SEI-87-TR-11, Characterizing the Software Process: A Maturity Framework [Humphrey 87b]. The SEI later refined and expanded this maturity framework when the Capability Maturity Model for Software was published (CMU/SEI-91-TR-24) [Pauk 91].
The DoD wanted a tool to help them evaluate the software processes of potential contractors. The result, SCE, assembled the positive attributes of other kinds of audits, some not necessarily from software, into a method with emphasis on observed capability. The SCE method is at its core a quality audit: its elements include interviews, document review, observations of operations, and the production of findings by comparing the data collected to a pre-established standard. By making the results of the evaluation a factor in source selection decisions, the government hopes to select better contractors and to encourage contractors to invest in software process improvement. The information SCE collects is timely, real, and based on current projects and the actions of engineering and managerial personnel.

SCE has evolved to be used in two ways:

1. **Source selection.** SCE has been used as a tool in source selection of software contractors. Most SCEs today are applied in this manner. The results of an SCE are used to characterize the software process-related risk of awarding a contract to a particular contractor.

2. **Contract monitoring.** SCE has been used in the monitoring of a contract after contract award by serving as the basis for an incentive-award fee decision or by focusing the monitoring based on observed contractor weaknesses. This use of SCE is still new but shows great promise for the future. It recognizes the long-term relationship of government and contractors and provides encouragement for investments in the software process by the contractor.

This paper focuses on the use of SCE in source selection because that is how SCE has primarily been applied to date. In source selection, the SCE goal is to determine the performance risk related to awarding a particular contract to a developer. By evaluating a contractor's software projects against the maturity framework, the SCE team can determine whether the contractor follows a stable, predictable software process (i.e., a process under control). Although a mature process does not guarantee a successful product, the likelihood of success should increase with a predictable and defined software process.

Since the development of the SCE in 1987, the method has gradually evolved as the SEI receives feedback from government and industry. This paper captures the experiences and lessons learned over years of observing SCE teams in action.
OVERVIEW OF THE SCE METHOD

Software Capability Evaluation involves more than a site visit to a contractor’s development facility. The activities surrounding the execution of an SCE to select a contractor or monitor an existing contract typically begin many months before an actual site visit. Likewise, the conclusion of an SCE may not come for one to two months after the final site visit.

The high-level activities performed to incorporate SCE into an acquisition process are presented in Figure 1. These activities include:

- Deciding to use an SCE in a contract award, which involves weighing the costs and benefits of an SCE. Both the contractor and the government incur significant costs to host or conduct an SCE.

- Planning the SCE, which includes deciding the role of the SCE in the contract, writing source selection documentation, and selecting and training the SCE team.

- Executing the SCE, which includes all the activities of the SCE team, such as pre-site visit planning, the site visit, and the presentation of findings to the contractor.

- Applying the SCE findings to the contract, where the source selection team evaluates each contractor's proposal and integrates the SCE findings into the source selection decision.

Figure 1 presents a hypothetical timeline showing each phase of SCE use. The precise amount of time for an event and whether it overlaps with another event, however, will be determined by the unique demands of each contract.
DECIDING TO USE AN SCE

The time to decide to use an SCE in a contract award is before the request for proposal (RFP) is released. This means that the government agency must understand the implications of using an SCE early in the acquisition process.²

Which contracts can benefit from using an SCE? If software development is anticipated to be a risk driver on a contract (i.e., a major contributor to performance risk, the risk of not meeting cost, schedule, or quality criteria), then SCE should certainly be considered. The government agency letting the contract needs to understand the factors that drive performance risk on the contract. These factors vary from one contract to the next but may include:

- **System size.** Large contracts are more likely to benefit from a focus on the software process, because larger systems are more difficult to manage, and the software process becomes more critical to the success of the contract. Size is often defined by contract length, lines of code, or number of people. For example, an estimated software cost exceeding 10 to 15 million dollars may indicate a contract that will benefit from the use of SCE.

- **System complexity.** Complex systems place a greater strain on software management because they require greater coordination among different parts of the development, contractor, and user organizations. If a contract requires several development teams at different sites or with multiple areas of expertise to coordinate with each other and with the customers, then managerial control of the development process will be critical to the success of the project.

- **Degree of precedence.** If similar systems have been built before, experience or knowledge may overcome deficiencies in process. But unprecedented systems or technologies generally require a more disciplined process to manage uncertain requirements and new issues.

The government agency must first analyze these and other relevant factors and identify all major sources of risk on a contract, such as technology, resources, staffing, and software process. Next, the agency must weigh these sources of risk to determine whether a contractor’s software process is likely to contribute to program risk. Finally, a qualitative cost-benefit analysis should be done to weigh the importance of software process with the cost and feasibility of using SCE.

² To help managers acquire an in-depth understanding of when to use an SCE, the SEI offers an SCE Overview Seminar. For more information, call (412) 268-5800.
The main costs to the government for an SCE include:

- **Time.** The SCE team usually spends at least two weeks preparing for and performing the SCE. That time includes preparing for each site visit, executing the SCEs, and producing the detailed report and findings for the government agency ordering the evaluation. The SCE team leader is needed for approximately six to nine months on a part-time basis prior to the site visit(s) to prepare all of the source selection materials that are affected by SCE, assemble the team, and schedule SCE training for any untrained team members.

- **Money.** The SCE team visits each contractor once during the source selection. Each trip costs the government agency approximately $10,000. The total cost increases significantly if a number of contractors are competing for a contract, or if complex contractor teaming arrangements require evaluations of several prime contractors and subcontractors. The training cost for SCE team training at the SEI must also be considered. Typically, tens of thousands of dollars may be spent on a single SCE by the government, although training is a one-time cost for teams that perform SCEs for multiple contracts.

Contractors also incur significant costs when hosting an SCE, often more than the costs to the government. The government agency needs to be sensitive to the difficulties of the contractor if it wants to encourage true process improvement rather than a bureaucratic response to a government requirement. For example, staff time that is spent to prepare for and participate in an SCE is a significant cost to the company, and it cannot be charged to a government contract. Contractor activities include identifying projects for SCE review, completing questionnaires and project profiles, preparing materials and coaching individuals to accommodate the site visit, and giving interviews. Many managers, staff members, and engineers from ongoing projects will be involved with each SCE. Because the SCE team evaluates current projects, activity on those projects is disrupted to accommodate the SCE. If a project is at a particularly critical stage, the impact may be greater. Several contractors have provided documentation showing a single SCE can cost $70,000 to $100,000.

The government office must apply common sense in weighing the costs and benefits of SCE. If the software process is determined not to be a major source of risk, the benefits of SCE may not be worth the investment. On the other hand, if the software process is a risk driver on the contract, then the cost and effort spent on SCE may be small compared to the benefits to be derived from awarding a contract to an organization with a more disciplined software process.
5 PLANNING AN SCE

Once a government agency has decided to use SCE for a procurement, it must plan the implementation. SCE results can be used in source selection in a variety of ways, based on the needs of the contract. Regardless of use, however, the government's intent to use SCE must be communicated to contractors as early as possible, preferably before the RFP is released.

Figure 2 is an example of an SCE planning timeline. Remember that the times are hypothetical and will vary with each contract award. The major activities in planning an SCE are described below.

5.1 Decide the Role of SCE in the Contract

Should SCE be used as an evaluation criterion or as a performance risk consideration? How will SCE results be factored into the contract decision? How will SCE and the SCE findings be used during contract monitoring (e.g., award fee plan)? If the results are to be used as an evaluation criterion in source selection, then their weighting or ranking must be detailed. The government should also have a vision for what software process capability it requires for the contract in question. While this vision is not now a mandatory activity in source selection, we recommend that it be translated into the key process areas that will form the basis of the evaluations.

Figure 2
Example Schedule of SCE Planning Stages
5.2 Conduct the Bidders' Conference

Communication between the government and the bidders is critical to the success of an SCE. Contractors should use the Bidders' Conference to learn about the contract, influence the source selection, voice concerns, and ask questions. Discussing the SCE at the Bidders' Conference will help assure that all parties know what to expect during the SCE. The government should make a detailed presentation of the method and those key process areas that will form the basis of the evaluations.\(^3\)

5.3 Write Source Selection Documentation

Appropriate wording should be included in the Commerce Business Daily announcement, the source selection plan (SSP), the RFP, and the evaluation standard to prepare the contractor for the evaluation. The RFP should also state what SCE-related materials the contractor must include in the proposal, preferably in a separate volume. Suggested wording for the RFP is provided through consultation with the SEI. Whatever is published in the RFP is exactly what the SCE team must abide by. In the past, unclear RFPs have resulted in SCE teams having problems providing clear guidance to obtain the right cross-section of projects and identify the proper development sites.

At this stage, the government must decide which key process areas within the maturity framework, such as Software Quality Assurance (SQA) and Configuration Management, will be evaluated at all contractor sites. This information is conveyed in the RFP.

5.4 Select and Train the SCE Team

The evaluation team should have four to six members, preferably five. This is small enough to reach consensus fairly easily, but large enough for diversity of experience and skills. Team members should have a cross-section of seven or more years of experience in software engineering, software management, and/or contract acquisition. The team should include at least one senior, knowledgeable individual and no more than two junior personnel. Team leaders also need to have excellent interpersonal skills.

The evaluation team must be trained to apply SCE effectively. The SEI course on Evaluation Team Training provides a common vision of what software process maturity means. Reliability and consistency can only be achieved if all team members operate by the same rules. Learning by experience and by role playing during training helps team members recognize their individual biases. Training also helps the team develop skills for interviewing, for conducting document review, and for identifying the process strengths and weaknesses that would need monitoring if a contractor wins the contract.

Training reinforces the concept of team building, leading to a team that acts from a common vision and respects one another's abilities. All team decisions must be made by consensus. Figure 3 shows a typical team decision-making process. A structured consensus-building process is critical to a successful SCE because the volume of information gathered by the team requires a group effort. No one person can effectively absorb and analyze so much information in an SCE, typically three days.

---

\(^3\) We recommend identifying the precise set of software practices within each key process area that will be validated at each site.
5.5 Request Candidate Projects

The contractor identifies about six candidate projects to be investigated by the SCE team. The projects should represent the contractor's typical software process implementation. A Maturity Questionnaire (MQ) and a project profile are to be filled out for each candidate project, typically by the software project managers. The MQ asks the bidder to respond to questions about a wide sample of software processes. The SCE team uses the responses as one source of input to identify issues during execution of the SCE. The project profiles are then used to evaluate the degree of similarity of the candidate projects to one another and to the proposed project. The government must ensure that contractors provide the right materials and project profiles.
6 EXECUTING THE SCE

This section describes the activities involved in an SCE, as shown in Figure 4. Although the site visit is the most visible part of the SCE, some activities precede it, and others follow it. All these activities contribute to a successful SCE—there are no short cuts.

6.1 Prepare for the Site Visit

To prepare for each site visit, the SCE team reviews the SCE-related materials submitted by the contractor. These materials should contain project profiles, Maturity Questionnaire responses, and any existing and documented software process improvement plans.

6.1.1 Select Projects

The SCE team selects three projects from the set of project profiles submitted by the contractor for evaluation, based on the similarity of the project to the contract, which can be determined by comparing project profiles with the contract profile. The profiles should include attributes such as application domain, language, and estimated size.

<table>
<thead>
<tr>
<th>Before Site Visit</th>
<th>During Site Visit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maturity</strong></td>
<td><strong>Paper</strong></td>
</tr>
<tr>
<td>Questionnaire</td>
<td><strong>Audit</strong></td>
</tr>
<tr>
<td><strong>responses</strong></td>
<td><strong>requests</strong></td>
</tr>
<tr>
<td><strong>Preparation</strong></td>
<td><strong>Evaluation</strong></td>
</tr>
<tr>
<td><strong>Identification</strong></td>
<td><strong>of</strong></td>
</tr>
<tr>
<td><strong>of</strong></td>
<td><strong>Software</strong></td>
</tr>
<tr>
<td><strong>Key Issues</strong></td>
<td><strong>Process</strong></td>
</tr>
<tr>
<td><strong>Project</strong></td>
<td><strong>maturity</strong></td>
</tr>
<tr>
<td><strong>profiles</strong></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5
Major Steps in an SCE
6.1.2 Analyze Maturity Questionnaire Response

The MQ responses provide one source of input to the team for focusing evaluation topics during the site visit. The analysis of the questionnaire responses includes determining where different projects answered the same question differently or where one project gave contradictory answers to related questions. These responses provide clues to the team to focus attention on those practices that may be weakly implemented.

One of the SCE team tasks during the site visit is to validate the responses given for the projects chosen by the team. The SCE does not simply look for “yes” or “no” answers to questions—it validates the existence of a procedure, its effective implementation, and whether the organization trains its staff to perform the procedure. This validation, however, is not the primary goal of the SCE. It is simply one method to help achieve the final goal—the identification of strengths, weaknesses, and improvement activities that influence process risk.

6.1.3 Analyze Software Process Improvement Plan

Although it is not required, some SCE teams request a Software Process Improvement Plan (SPIP) for the development site being evaluated. The SPIP is the document, in the contractor’s format, that contains its software process improvement road map. The SPIP provides a useful way of gaining insight into the strength of a contractor’s improvement activities, which may mitigate the performance risk for a contractor, especially in process areas that are determined to be weak in the SCE findings.

The SCE team should not ask for Software Process Assessment (SPA) results, either in the RFP or while conducting the on-site evaluation. SPAs and SCEs are different and have different purposes. Assessments seek to facilitate process improvement, while SCEs seek to determine process capability relative to a specific contract effort.

6.1.4 Prepare for Interviews

The SCE team must prepare for interviews before the site visit. To prepare, the team must:

- Select personnel to interview. The team should select personnel by function (e.g., SQA manager) rather than by title.
- Plan an interview schedule with the contractor’s SCE coordinator. The team should be sensitive to conflicting commitments on the contractor’s part and maintain some flexibility in its schedule to account for last-minute changes.
- Prepare interview “scripts” for each interview. The team should identify questions or topics to discuss with each interviewee. At a minimum, each team should have scripts prepared for the senior manager, project managers, software managers, and various functional representatives. Here, too, the team should be flexible and anticipate changes to the scripts as the site visit progresses.

6.1.5 Coordinate the Site Visit

The SCE team should be responsible for coordinating the site visit with the contractor concerning the facilities and schedules needed to effectively conduct the visit (e.g., meeting...
rooms, phone access, and notification of the presentation by the government, which begins the SCE). If possible, to be fair and consistent, the SCE team should give each contractor exactly the same amount of time between notification of the three or four projects being examined and the start of the SCE site visit.

6.2 Conduct the Site Visit

The site visit is the heart of the evaluation method. The goal of the site visit is to understand the processes that are followed by the projects. A great plan or a superb standard is of little value if it is not followed. The evaluation team will reach an understanding of an organization's software process by conducting interviews and reviewing working-level documentation and organizational policies and procedures.

6.2.1 Initial Presentation

An initial presentation enables both the SCE team and the contractor to establish expectations and ground rules for the site visit. All contractor key managers, functional representatives, and interviewees should attend. The SCE team should discuss the SCE process and outline the various activities, for example, how the interviews work, the likely individuals by job type to be interviewed, the types of questions to be asked, and a sample of the documentation to be examined.

The contractor should describe its organization, discuss its software process, how it manages the implementation of its software process, and its software process improvement activities. This is not a marketing pitch but an overview of contractor operations and a status of each project to help the SCE team understand how the site is organized. The total presentation should be limited to two hours.

6.2.2 Team Orientation

The SCE team generally begins the site visit by reviewing organizational policies, standards, procedures, and project status information. The team needs to spend this time getting oriented to the contractor to maximize each interview. Every contractor organizes and manages its people differently, has differences in its software process, and uses unique vocabulary and acronyms to describe its process. Without a good understanding of the contractor's organization, software process, and project status, the team could interview the wrong people or ask the wrong questions. The team can minimize these possibilities by spending a little time up front developing a team perspective.

6.2.3 Exploratory Interviews

Next, the team conducts exploratory interviews to understand the implementation of the organizational processes. The team asks managers and project personnel to explain how they actually do their job. For example, the following questions could be asked about activity X:

- Will you describe how you do X?
- Who is responsible for doing X?
- What are the outputs of X?
- How do managers control X?
- Can you show us audit trail information on the execution of X?
Generally, senior personnel are interviewed first, followed by program managers, software managers, project leaders, and staff, although this strategy is not required. Most of the time should be spent on project personnel rather than on top- or mid-level management. The questions asked at all levels need to be commensurate with the level of the person being interviewed. For example, questions to senior managers should be related to issues of policy and how they manage the performance of their projects.

After the managers, the SCE team continues down the organization interviewing various functional representatives and first-line software development managers. Who is interviewed and for how long is largely dictated by what the team has learned to this point. Naturally, some of the software quality assurance, software configuration management, and software engineering process group (SEPG) personnel, both managers and staff, need to be interviewed—however, findings accumulated to this point should dictate the degree and amount of time devoted to each of these groups.

Important groups of people to interview are the actual software developers, testers, or any other personnel pertinent to an issue being evaluated. At this level, the team seeks to validate the implementation of a practice or process. Are the processes identified in the plans and procedures really being used? Can the individual who actually uses the process articulate what is done? These are some of the questions the SCE team needs to think about as it develops its interview strategy.

6.2.4 Interview Style

The goal of all SCE teams during an interview is to collect information while making the interviewee as comfortable as possible. To this end, SCE training classes prepare teams to achieve their purpose while putting the interviewee at ease, for example:

- **Introduce all team members.**
  Explain the purpose and nature of the interview at the start. All interviews are confidential to the SCE team—no names are released to the contractor or to the government agency.

- **Use only one primary interviewer.**
  Avoid the “many-against-one” perception. The rest of the team primarily listens to responses, and can ask follow-up questions at the end of the interview.

- **Use polite interruptions.**
  Keep the interview focused. The team has limited time and needs to maximize the usefulness of the interview. At the same time, the team must provide the interviewee adequate chances to present appropriate information to answer the questions.

- **Ask open-ended questions.**
  Simple “yes” or “no” questions do not help validate responses. The team should ask the interviewee to describe procedures, policies, and specific events (e.g., meetings) that can be examined by the team during subsequent document review.

- **Allow the interviewee to ask questions or clarify responses.**
  Sometimes the interviewee may misunderstand the intent of a question, which can frustrate the team’s attempts to gather the right information. Interviewees can use this opportunity to ensure that the SCE team clearly understands their responses.

Unless teams encounter overt contradictory information, SCE teams are taught to accept what they hear during the interviews as fact. We emphasize overt because many times what one individual on an SCE team may perceive as contradictory information is just a difference in perception within an organization on how some facet of the software process is implemented. The team must avoid becoming argumentative or confrontational during the interviews.
The SCE team should remain focused on its objective—to collect information. It is not there to debate any facet of the maturity model, discuss what a project should do programmatically, engage in discussions on the latest software tool, or advise a manager or staff member on what they should be doing with a particular piece of technology. The team should pay attention, thank interviewees for their responses, and remove any individual from the room and maybe even from the team who cannot act professionally. While an SCE team is evaluating an organization, it, too, is being evaluated by that organization, and any misbehavior by a team member could reflect badly on the team, the program office, and the government.

6.2.5 Document Review

The purpose of document review is to validate the existence and consistent implementation of software practices that have been described in organizational procedures or interviews. The team conducts document review throughout the site visit as needed. An important lesson in reviewing documents is that there are significantly more documents than there are people and time available to review them. As a result, the team leader needs to ensure that the team is examining only essential pieces of documentation. In most cases, document review really consists of spot checking or verifying that a particular practice is really implemented. To the greatest extent possible, no finding should be validated without documentation to support it.

The review of audit trail information or low-level documentation representing the implementation of a particular practice is an important activity for the SCE team. The team determines whether the document trail demonstrates consistent implementation of activity X. Few processes are implemented without a trail of some sort, such as meeting minutes, action items, and reports. The formality of the documents may differ from project to project, but the trail should exist and should allow the evaluation team to investigate the actual, not just the planned, process.

6.2.6 Engineering Floor Walk-Through

Walk-throughs are used by some SCE teams to escape the somewhat artificial interview and document-review setting. Walk-throughs are used extensively in a number of financial and quality audits performed by a variety of organizations. The team goes out to look firsthand at the implementation of the process on the work floor. For example, the team may: review software development folders and software quality assurance records, observe ongoing testing, attend a software configuration control board meeting, or talk informally with some of the developers about the software process and the support they receive to perform within that process.

There is no precise agenda for this activity; it is guided by what the team needs to learn firsthand. If the team does decide to conduct this activity, the organization should be informed as early as possible (prior to the site visit) to prevent undue disruption of meeting and work.

6.2.7 Consolidation Interviews

The team may conduct consolidation interviews to resolve any lingering question and request additional documentation as part of a consolidation interview to substantiate
The team tries to get at least two sources of information, written or verbal, to corroborate a finding that it thinks may be significant.

Some teams have used consolidation interviews to informally communicate preliminary findings to the contractor as an additional opportunity to the contractor to provide appropriate data. An easy way of accomplishing this is to say: "The evidence we've gathered indicates that X is not in place organizationally. Are we missing something? Do we need to talk to anyone else?" This is a valuable tactic for the team to employ because it generates data that either reinforce a finding or produce evidence to the contrary.

6.2.8 Professional Judgment and Validation

As the team gets closer to developing its findings, it must apply its professional judgment to reach consensus about its observations. The team may judge two aspects of a process: whether the contractor is following a reasonable process, and whether the process is effectively implemented. SCE team training emphasizes the need to tailor expectations to the specific circumstances of a given project. Few issues the SCE team faces are cut-and-dry decisions. For example, teams will see older projects managed differently from newer ones.

A reasonable process is one that helps the project achieve the goals of the particular key process area. The team does not judge the reasonableness of every process it evaluates—a process is assumed reasonable unless concrete evidence indicates otherwise. Nor does the team ask, "Is this process implemented the way we think is best?" Many different implementations of a process are possible, and the team needs to be flexible during its evaluation.

The team also validates that a process demonstrates elements that can be reasonably expected in an institutionalized process. These elements include policies and procedures to define the process, training programs to communicate the process, the implementation track record, defined roles and responsibilities, and adequate resources. Several examples of what constitutes an effective implementation of a practice are presented during SCE training.

The needs of the individual contract determine the exact definition of a validated practice. For example, within one source selection, the process rigor required may be great; therefore, what constitutes validation by the team for that source selection may be substantially more rigorous than what constitutes validation for another source selection.

6.2.9 Determination of Findings

The result of the site visit consists of evaluation findings that describe the strengths, weaknesses, and improvement activities of the contractor. The team generates findings for each key process area (KPA) that is evaluated from the maturity framework. It focuses on those processes that it identified before the site visit as most important to risk on the contract. The team should state general findings that are issues for the entire organization, and avoid absolute statements by using phrases such as "lack of evidence." The team should always be prepared to back up statements with facts but without attribution of the sources of the data.

One concern in performing evaluations is whether an organization might try to bias an evaluation, for example, to answer questions "yes" without actually having implemented the practices underlying the questions. Although
this is a legitimate concern, it has not been a major problem. A trained team identifies any 
"holes" in the implemented process as an integral part of the site visit.

6.2.10 Presentation of Findings

Where possible, the evaluation team should present its findings to the organization at the end of the site visit. For contracts that do not allow SCE teams to present findings before contract award, the findings should be presented at the earliest possible time. The motivation for presentation is to encourage the contractor to use feedback gained from the SCE. Contractors gain a customer's perspective of their process, which provides input for their process improvement efforts. The debriefing also gives contractors an opportunity to clarify misunderstandings, but should not degenerate into an argument over the findings.

6.2.11 Completing the Site Visit Report

At the conclusion of each site visit, the SCE team should document for the government agency its findings in a report or as an addendum to the findings slides that may previously have been presented to the contractor. That report should capture the essential facts supporting each finding, but not identify any names as sources of that information. The team needs to prepare this report in such a manner that it is willing to submit it to authorities outside the source selection team in the event of a protest. It is important that the site visit report be prepared within a few days of the site visit.

---

4 Different services, commands, and government agencies have interpreted the federal acquisition regulations (FARs) quite differently. For some contracts, no one on the evaluation team was involved in the contract except during the evaluation, to prevent accidental transfer of contract-sensitive information to a bidder. For other contracts, the evaluation team consisted of the Source Selection Evaluation Board. For still other contracts, any feedback to the contractor was forbidden, some allowed a debriefing to the contractor at the end of the site visit, and others debriefed following contract award. The legal and procurement concerns center on the issue of discussions (i.e., does an outbriefing constitute discussions?). The information SCE captures is prior and current performance information from existing projects; it is not an evaluation of any facet of a proposal (however, the source selection organization can and will compare the SCE results to what is being offered in the proposal). When properly handled, any of these approaches conforms to the FARs.
In source selection, the SCE findings are integrated into a statement of risk, a numerical rating, color code, or other kind of rating scheme that compares the findings to the pre-established evaluation standard, and produces a rating useful to the respective source selection. For example, the Air Force uses a color code-based scoring scheme in which evaluation criteria or “items” are coded either blue, green, yellow, or red, with blue being the highest. A risk factor along with a statement of risk is attached to each color score.

To arrive at the colors and risks, the SCE team leader, along with the source selection organization (the group that prepares the ratings), apply each set of findings to the pre-established evaluation standard. At this stage, contractors are not compared to one another. The colors and risks assigned to contractors are based strictly on their performance, not on how well they did compared to the others. Figure 5 provides an example of this process.

For each contractor, the technical (color) rating, strengths and weaknesses, risk, and a statement explaining the basis for the risk are identified and placed next to those for the other

<table>
<thead>
<tr>
<th>Item: T-3 Software Engineering Capability</th>
<th>Contractor A</th>
<th>Contractor B</th>
<th>Contractor C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>STRENGTH</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer Reviews</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standards and Procedures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WEAKNESS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software Engineering Process Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RISK</strong></td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>They are very strong technically and committed to developing quality software using a continuously improving development process</td>
<td>Because of the large disparity between our findings and their internally assessed maturity level, we find it highly questionable whether they have a a genuine software process improvement program.</td>
<td>They are strong technically and committed to developing quality software using a continuously improving development process; however, they are still at an initial maturity level.</td>
<td></td>
</tr>
</tbody>
</table>

$L = \text{low risk}$  
$H = \text{high risk}$
contractors so that the source selection authority (the individuals who decide the contract award) and advisors may compare and discuss them during a presentation.

Figure 5 also illustrates one way the Air Force has assigned risk in the past: Contractor B has a high risk (designated by an H) and Contractor C has a low risk (designated by an L); yet Contractor C has four weak key process areas, and Contractor B has only two. How could this happen? In this example, risk was based upon the consistency of the organization’s process improvement program to the SCE findings. Because it was stated clearly in the RFP for this contract that an organization could be at a low maturity level and still be awarded the contract, this source selection organization defined risk as a measure of the integrity of an organization’s software process improvement program, regardless of its stated maturity level rating in the Software Process Improvement Plan (SPIP). The SCE team found that Contractor B’s SPIP indicated an unrealistically high level of maturity, relative to the SCE findings. On the other hand, Contractor C’s SPIP indicated a low level of maturity, yet it attempted to apply its best practices to all new programs. The SCE findings confirmed this, which resulted in the assignment of a lower risk to Contractor C than to Contractor B.

This example is not the only way or even the recommended way to assign risk. An acquisition should tailor the risk assignment to the specific needs of that contract. Additionally, this example used maturity levels in the formulation of the risk, contradicting our earlier advice against the use of maturity levels. That was a decision made by the government agency, but it did not use the maturity levels as filters in making the source selection decision. Furthermore, it used a loose interpretation of maturity levels, basing them on what key process areas were strong, acceptable, or weak rather than on a maturity questionnaire analysis.

The last step in applying the SCE result is integrating the SCE technical rating and risk factor with those of the other technical items to produce a technical area summary, shown in the example in Figure 6. Here the first three items were assigned a color code and risk while the last two items, examinations on
multi-level security and Ada expertise, received a rating of “AA” for above average, “A” for average, or “B” for below average. The technical ratings and risk factors were then integrated into a summary rating for the technical area.

There is no standard rule across contracts for integrating color codes into area summaries other than one pertaining to the red color code: if any item in an area is assigned the red color code, then the entire area must also be red. The source selection organization integrates the color codes and risk factors into area summaries based upon its expertise and recommends them to the source selection authority. The procurement and legal advisors ensure that the colors have been integrated in a consistent manner for all contractors.

After the evaluation is performed, we encourage the SCE evaluation team and evaluated organizations to provide feedback to the SEI on what worked and what needs improvement to help the SEI improve the SCE method and team training.
8

COMPARING SCE AND SPA

Much attention has been drawn to the fact that SCEs and Software Process Assessments (SPA) often do not get the same results. By far the most important lesson learned confirmed what we suspected: comparing the results of evaluations to assessments is like comparing apples to oranges, especially when viewed through the maturity-level lens.

While there are several notable exceptions, unless a large number of variables are controlled, it is inconclusive to draw conclusions solely from the differences in maturity level findings produced by each method. SCEs and SPAs differ in many respects, including motivation for the method, objective, ownership of results, and outcome. These characteristics lead to several differences in the conduct of the methods. Some of the major differences between evaluations and assessments are described below.

8.1 Project Selection

SCE teams select projects to evaluate based on their similarity to the contract for which the SCE is being conducted. SPA teams selects projects that are representative of an organization or site as a whole.

8.2 Basis of Evaluation

The Maturity Questionnaire from CMU/SEI-87-TR-23 [Humphrey 87b] is only a sample of very comprehensive key process areas. Because the underlying processes were not clearly defined in TR-23, early SCE and SPA teams were evaluating organizations against the questionnaire without a clear understanding of what constituted a "yes" or "no" response to a particular question. Sometimes this resulted in an emphasis on paper processes rather than process substance. This issue is not yet resolved, but the software community in general now has a greater appreciation for the software processes (i.e., the maturity framework) underlying the Maturity Questionnaire and how they can be implemented in an organization.

8.3 Investigative Methods

SCE relies primarily on one-on-many interviews and document review to gather objective evidence for a contractor selection decision. SPA, on the other hand, takes a collaborative approach, relying primarily on one-on-one interviews and group discussions among organizational personnel to encourage an atmosphere of process improvement. Each method was developed for a unique purpose, and these differences can influence results due to the group dynamics involved.

8.4 Lack of Familiarity vs. Familiarity with the Development Organization

An SCE team lacks detailed knowledge of how an organization operates—it searches for every piece of information to make decisions about the software process. In three days, this is a difficult task, and requires an experienced and trained team that is conscientious about not jumping to hasty conclusions without objective supporting evidence.

A SPA team, on the other hand, may be hampered by too much familiarity with an organization. It may make assumptions about "how things really work" without applying the objectivity expected of an SCE team. Some industry representatives from SEPGs have called for more rigor in the way SPA teams develop findings.
9

ISSUES FACING THE SCE METHOD

The following are some of the most important issues facing the SCE method today.

9.1 Refresher Training

The many changes to the SCE method, which we discuss in the next section, need to be communicated to early trainees to ensure that all SCE teams are working from a consistent baseline. The SEI is currently baselining the method taught in 1992, and plans to release a major revision to SCE team training in 1993. Refresher training will be planned at that time.

9.2 Importance of Teamwork

SCE teams can only succeed if they operate as a team and not as individuals. Each individual's role on the team must be clearly defined, and the team must be committed to a true consensus-building process. The SEI has addressed this issue in part by extending the SCE training course an additional day to allow more time for case study exercises. Course revisions to be released during 1993 will further extend the training to five full days.

9.3 Feedback from Industry

The SEI recently gathered valuable feedback about the SCE method at forums such as the SCE Users Workshop held in July 1992. In addition, the recently formed SCE Advisory Board, composed of industry and government representatives, will play a pivotal role in determining the strategic direction of the SCE method. Some important considerations from industry are presented below.

9.3.1 Knowledge of SCE

Industry's most important issue seems to be the lack of knowledge about the SCE method. Companies require this knowledge so that they can adequately prepare for an SCE. The SCE Users Workshop helped in this regard. The strong demand from the government for SCR training has precluded the SEI from admitting industry teams to the SCR training course. In 1992, the SEI began offering a one-day SCR Overview Seminar that is open to both industry and government.

9.3.2 Differences Between SCE and SPA

Industry is concerned about the differences between SCE and SPA results, which may have an impact on senior management commitment to SPA and process improvement efforts. Some reasons for these differences were discussed in the previous section.
9.3.3 Cost of Repeated SCEs

As we discussed earlier, a single SCE can cost a contractor $100,000. Some of the larger contractors have experienced several SCEs in the space of a few months, which is expensive and disruptive to their operations. They would prefer a method whose results can be reused on several contracts. While this is not feasible with the current SCE method, in the future, the SCE method may address this need more directly.

9.3.4 Private Interviews

A common concern from industry is the use of private, confidential interviews during an SCE without a company representative present. Government agencies counter that the confidentiality of the interview increases the likelihood of more authentic responses from interviewees. Both positions have merit. Government audits typically afford the contractor the opportunity to have some type of representation in the room during interviews. However, Software Capability Evaluations make private interviews a stipulation identified in the RFP. This is an issue that will be discussed in future SCE workshops, but no change in the SCE approach has been incorporated into the method at this time.

9.3.5 Diverse Organizational Structures

Increasingly, government contractors use complex teaming arrangements, subcontractors, and multiple sites to pool expertise. Often it is difficult to identify the proper development site because of multiple contractor team members who operate from geographically distributed locations. This issue has not been resolved and will need to be revisited as SCE use expands throughout the community.

10 THE EVOLVING SCE METHOD

The maturity model, questionnaire, and SCE and SFA methods are documented [Humphrey 87a], [Humphrey 88], [Paulk 91], [Weber 91] and updated as experience is gained with their use. These products are currently being revised in light of both the SEI experience and feedback from users.

When field-test use of the evaluation method began in 1987, evaluations were not performed according to the process described here. Today's version of the SCE method has evolved significantly from CMU/SEI-87-TR-23, the first version [Humphrey 87b]. While the major steps, such as interviews and document review, have remained the same, other aspects have changed significantly. Some concepts that are taught in SCE training today that were not taught in 1987 are presented below.

10.1 Elimination of Maturity-Level Scoring

The original SCE method included a mechanism to determine a contractor's maturity level score. Feedback from SCE teams demonstrated that maturity level scores are not useful indicators of actual process capability because they are abstractions of the underlying process strengths and weaknesses observed by the SCE team. Maturity levels are still useful as descriptors or goals for process improvement, but they are not used in the current SCE method, except to identify the scope of the SCE (i.e., which key process areas will be evaluated by the SCE team).
10.2 Shift from Maturity Questionnaire to Key Process Areas

The original method relied on the SEI Maturity Questionnaire to “sample” broad key process areas in a contractor’s software process. Information was collected to verify that an organization’s responses to the questionnaire were based upon actual practice, and then to determine the organization’s software process capability by scoring the validated responses. This was the “question-based” SCE method. However, many industry and government people thought that the resulting sample was an inadequate representation of a very complex process. By 1989, SCE had shifted the emphasis away from the MQ to a more direct evaluation based on the underlying key process areas. The current version of SCE uses the MQ only as one tool to focus the team’s evaluation of key process areas.

10.4 Evaluation of Contract-Specific Risk

By emphasizing the Maturity Questionnaire in the past, SCE teams looked at essentially the same software practices for all uses of the SCE method. That is, no thought was given to which of the many practices embodied in a given key process area might be more relevant to the contract at hand than those sampled by the questionnaire. By emphasizing the maturity framework over the questionnaire, SCE teams are in a better position to review what is important for a given application of the method. Different teams may look at different practices within each KPA, based on the needs of their contract. The individual strengths, weaknesses, and improvement activities observed in these practices form a picture of the process risk to the contract.

10.3 Definitions of Key Process Areas

Due to the shift from the MQ to key process areas, SCE teams no longer evaluate information related to specific sample software practices (i.e., MQ questions). Rather, they make judgments about each key process area by observing and collecting a broad range of information. To help SCE teams evaluate key process areas effectively, the SEI has defined extensions to the maturity framework. These extensions define each key process area as a set of subprocess areas that are smaller in scope to enable a more focused evaluation. In addition, a set of features common to all subprocess areas helps generate criteria to determine if a subprocess area is satisfied by the contractor’s process (the elements include policies, procedures, training programs, resources, measurements, roles and responsibilities, and implementation mechanisms).
11
SUMMARY AND
CONCLUSIONS

The use of SCE is expanding rapidly because government agencies have found that the method enables them to gain insight into a contractor's software development capability by a means not offered through any other tool. Software capability evaluation can be a powerful tool for government agencies, but requires training, time, and dedication if those agencies are to use it effectively.

The process of incorporating software capability evaluations into the source selection process must begin months before any site visits. A government agency must weigh system size, complexity, application domain, target software process capability, and a multitude of other program-unique issues as it decides and then plans the implementation of SCE into its source selection process. The process ends with the site visits, translation of findings into a risk assessment, and the resulting contract award.

The SCE method has evolved over the years as industry and government have transferred their lessons learned to the SEI. The SEI continues to evolve the SCE method based on feedback from SCE teams and industry.
REFERENCES


Joseph Besselman is a resident affiliate in the SEI Process Program from the Air Force Materiel Command and a doctoral candidate at Carnegie Mellon University.

Paul Byrnes is the project leader of the Software Capability Evaluation Project at the SEI.

Cathy Lin was a resident affiliate from the Air Force. While at the SEI, she worked on the Domain Analysis Project, which has now been incorporated into the Application of Software Models Project.

Mark Paulk is the project leader for the Capability Maturity Model (CMM) Project. His work focuses on identifying best software engineering practices and documenting them in the CMM and CMM-based products.

Rajesh Puranik is a member of the technical staff working on the Software Capability Evaluation Project.