SOFTWARE PROCESS IMPROVEMENT AND PRACTICE
Softw. Process Improve. Pract. 2003; **8**: 145–156 (DOI: 10.1002/spip.175)

Practical Recommendations to Minimize Software Capability Evaluation Risks

Practice Section

Hossein Saiedian*,†
Electrical Engineering and Computer Science, University of Kansas, Lawrence, Kansas 66045, USA

Many organizations, and, in particular, the US Department of Defense, find the Capability Maturity Model's (CMM) software capability evaluation (SCE) to be a useful tool for evaluating the maturity and capability of a number of competing software contractors and thus selecting a prospective contractor. The evaluation commonly takes place during the acquisition phase of a software project in order to focus on identifying and reducing risks associated with quality, scheduling, and budget. When an unbiased evaluation is performed, the information revealed can help determine potential risks in selecting a particular candidate contractor. However, there are many ways to manipulate the outcome of the SCE by the evaluation team or the candidate contractor. It is important that the potential flaws and shortcomings in the SCE be identified and eliminated to ensure a viable evaluation. This article briefly discusses the concerns associated with and flaws in conducting SCEs and proposes a number of practical recommendations to eliminate or substantially reduce them. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: capability maturity model (CMM); software capability evaluation (SCE); quality software development; risk identification and management; schedule and budget

1. INTRODUCTION

When a software acquisition agency, for example, the US Department of Defense (DoD) issues a request for proposal (RFP) for a new software application, it is possible that many companies will respond with an offer. Since most of the software that the US DoD or any large organization needs may have many performance, stability, security, and similar requirements, it is important that a competent ('mature') company is contracted. This was the driving force behind the creation of the SEI's software capability evaluation (SCE). The SCE is intended to serve as a formal process that can determine the capabilities and associated risks involved with a particular contractor. Interviews and research are conducted by an evaluation team, which later reports its findings to the acquisition team, which in turn submits its findings to an acquisition organization (e.g. US DoD's 'selection advisory council') for final selection. The result of SCEs help US DoD[1] to select a candidate contractor; thus, it is very important that the results obtained from an SCE effort are genuine and unbiased.

The unfortunate truth is that this process is sometimes taken advantage of by both the evaluation

* Correspondence to: Hossein Saiedian, Electrical Engineering and Computer Science, University of Kansas, Lawrence, Kansas 66045, USA
† E-mail: saiedian@eecs.ku.edu

[1] In this article, we make many references to the US DoD for several reasons: (1) it is one of the biggest, if not the biggest purchasers of software, (2) DoD projects normally risk budget and scheduling concerns, and (3) the Capability Maturity Model and SCE were primarily developed for DoD purposes, although both have been used in other sectors, though not as often.

teams (to consciously or perhaps unconsciously favor a certain contractor) or by the candidate contractors (by presenting a false 'capable' and 'mature' image). Furthermore, some of the SCE requests are made with no proper planning and thus their results are not used at all. A comprehensive study was performed on concerns and the shortcomings in using the SCE in (O'Connell and Saiedian 1995):

1. *Acquisition organization side*
   - *SCE team selection policies*: As documented in (O'Connell and Saiedian 1995), the team selection policies and guidelines are loosely applied and thus the evaluation teams frequently lack individuals with sufficient depth in software engineering to evaluate a contractor.
   - *Too broad methods*: A properly conducted SCE may yield many interpretable results. It is important to understand the specific results and to make sure that they are not being applied frivolously.
   - *Selection of Evaluation Tools*: There are a large number of software and contractor evaluation tools available (e.g., SCE, ISO-9001, Mil-Std-498, ISO/IEC-12 207, IEEE J-Std-016). Often, government agencies tend to apply evaluation tools without discrimination with no well-defined and stated plans.
   - *Wasted resources*: A single SCE performed often costs up to $10 000 and one week time. Considering that the US DoD must evaluate multiple candidate contractors, SCEs can cost tens of thousands of dollars.
   - *Schedule economics*: An SCE team typically has one week to evaluate a candidate contractor. During that time, the team must review all project artifacts and conduct interviews, compile findings, determine contractor's CMM rating, and present their findings to the DoD's selection advisory council. It is difficult to conduct an SCE in a short time and still produce quality results.
   - *Using SCE results:* In some situations, the SCE results may be taken out of context, sometimes with too much importance, but sometimes the low ratings (and thus associated risks) are overlooked in favor of political and monetary concerns. It is also possible that the selected contractors (based on their high SCE ranking) are not as capable as their initial SCE indicated.
   - *Judgment calls:* SCE teams sometimes are required to make judgment calls regarding a contractor's fulfillment of a key process area (KPA). There are some situations in which a judgment call may result in an alternative process not being properly identified to fulfill the KPA, or an acknowledgment of fulfillment when the process is actually not being carried out.
2. *On the contractor side*
   - *Intentional vagueness*: The contractor may embellish on maturity questions, which can result in the SCE team accepting whatever answer was given instead of properly investigating.
   - *Intentional detail*: Contractors may overwhelm an SCE team by giving them too much or extraneous information during an evaluation. The SCE team will then not have time to process all of the information and may give credit where it is not deserved.
   - *Inappropriate sample projects and personnel*: The contractor may provide nonexemplary projects or personnel in order to make them score higher on the SCE.
   - *Coaching*: Employers may coach potential interviewees so that they will give more satisfactory answers during an SCE.

When analyzing these concerns, it is important to consider ways in which they can be resolved without any major alterations made to the SCE process itself. Many times the problems can be solved by making sure that the SCE is performed by properly trained personnel. This article will present several solutions that will cover these problem areas. The first set of solutions are to be applied before the evaluation, the second set to be applied during the evaluation, and the last solution to be performed after the contract is awarded, as shown in Figure 1. Solution areas are represented in the next section.

2. PRACTICAL RECOMMENDATIONS TO IMPROVE SCE

In this section, we provide our recommendations to make SCEs more viable, effective, and less risky. Our recommendations are based primarily on personal experiences and also on communicating with experts in the field.

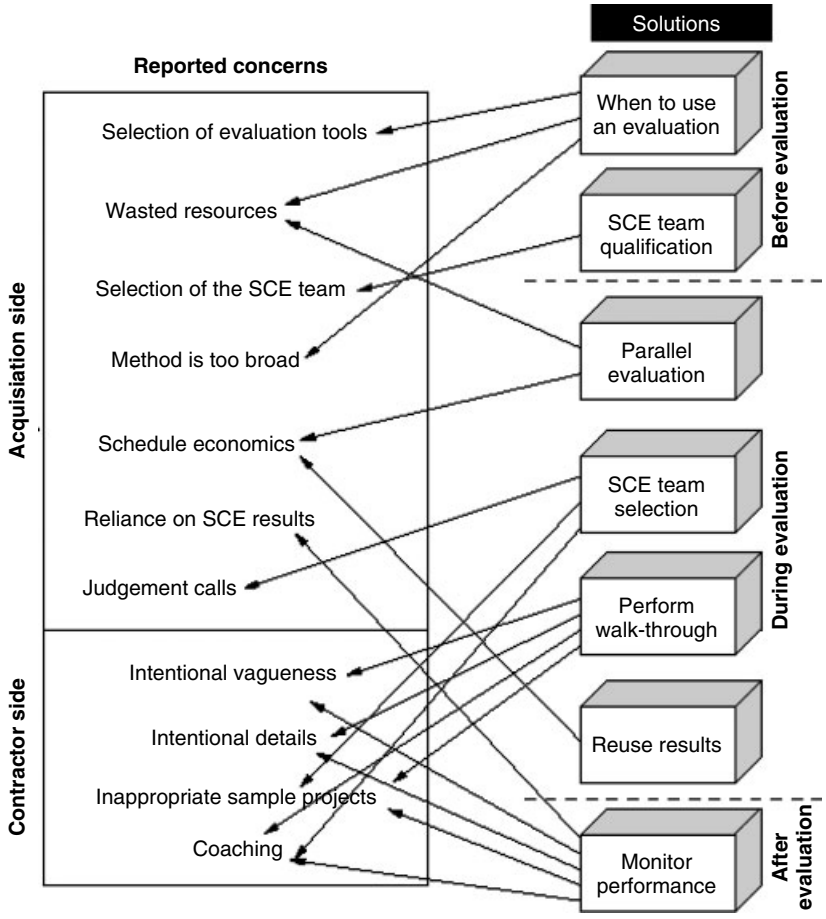

Figure 1. Documented SCE problems and solution areas

2.1. Base the Decision to Use the SCE on Importance of Maturity and Potential Risks

An evaluation should be conducted when the work to be performed requires a contractor with a mature software process or represents a significant degree of risk.

Owing to the amount of money and time it takes to complete an SCE, it is necessary to have certain criteria to consider before performing an SCE. Inconsistent usage of the SCE can lead to high costs and lost time, among other wasted resources. There needs to be a form of initial filtering of projects that eliminates the cost of unneeded SCEs being performed.

An example of an unneeded SCE is provided by O'Connell and Saiedian (O'Connell and Saiedian 1995). In 1995, an SCE was performed for an Air Force contract. The winning contractor was a CMM Level 3, but they were integrated with the Air Force software-development team, which only ranked at CMM Level 1. The Air Force's development practices were forced on the contractor. Since these processes were less mature, the integrated team ultimately produced a low-quality product. After a year was spent correcting problems, the US DoD approved a commercial product to replace the failed attempt. In this scenario, the SCE was unneeded since process maturity was not seemingly important.

In order to ensure that an SCE is performed consistently and in the proper situation, guidelines should exist that establish when the investment in performing the SCE will pay off. The first guideline concerns maturity of an organization. An SCE should only be performed if the maturity of the software processes of a candidate contractor is important. In the above example, since the contract was awarded to a contractor who was forced to

adopt the process of a less mature organization, maturity was not really an issue.

The SCE technique provides an all-encompassing view of an organization's maturity. It evaluates the processes used by the contractor on projects that are similar to the one being bid. It is based on the CMM and addresses no specific technical expertise. Overall, use of the SCE anticipates that a well-defined and measured process will be self-correcting. As long as there is sufficient insight into the results of the process, a contractor will be able to correctly adapt processes for an individual software effort in order to achieve quality and mitigate risks.

When is process maturity a concern? An example would be any large software project that is mission critical and will require maintenance over a long period of time. In this scenario, the CMM SCE will provide insight into the ability of the organization to implement repeatable, well-defined, and improving software processes (Paulk 1999). The same would apply for rapid prototyping projects, sustaining engineering projects, and research and development organizations. In each one of these situations, the project will benefit from code reviews, well-defined tests, and an ever-evolving software process (Humphrey 1988). It is for these reasons that this type of contract should be considered in the application of an evaluation technique. One contract does not necessarily consist of only one development effort. A contract can include development and maintenance of multiple software products. Here, a contractor's capability and maturity may be more critical because their processes and techniques will be used to create and maintain an entire suite of systems. This suite might comprise the entire mission and livelihood of the customer. In this case, the customer's total dependence on the contractor's performance warrants a formal evaluation.

Determining process maturity is not necessarily the only reason making the SCE worthwhile. When a significant degree of risk exists around the project, such as budget, size, and complexity, it may also be important to perform the SCE.

We need to define the term 'degree of risk' with regard to software projects. According to the CMM, the degree of risk in software system corresponds to the size, complexity, and uniqueness of the end product (Besselman *et al*. 1993). System size can be measured by lines of code, development schedule, or monetary cost. System complexity refers to the technical requirements of the system. This might include hard-real time, security restrictions, number of external interfaces, or required system architecture. System uniqueness describes whether a similar product exists and the likelihood that the software engineers will have references or previous experiences upon which to base their efforts. Each one of these components should be used in deciding if the investment in performing an SCE is beneficial.

The US DoD lists five criteria in its definition of software risk:

1. The software product is critical to the system's mission accomplishment. Software is almost always an essential piece of US DoD systems.
2. The software constitutes a major portion of the overall development effort.
3. A primary component of the system, including its software functionality is considered to be unprecedented. This is true of most US DoD software applications, and that is why commercial products are rarely an option.
4. The software-development cost exceeds $5 million.
5. Any software developed during a prototyping or validation phase will be utilized during a follow-on contract.

According to a 1994 policy issued by the US Deputy Assistant Secretary of the Air Force for acquisition, a software evaluation is required if even one of these risk criteria is met. Beyond the software risks imposed by any large project, there are additional risks that apply to the US DoD's acquisition process that may lead to some of the problems discussed earlier in this section. Without proper measures in place, it is possible that a lack of oversight on behalf of the US DoD may result in contractor biases. The degree of oversight describes how closely the government customer will interact with the contractor during the software-development effort. It is directly affected by the high personnel turnover rate within the US DoD. It is not uncommon for US DoD program managers and engineers to be 'promoted out' of an acquisition supervisory position or be relocated to a different organization. Thus, program managers are not held accountable for their programs (Saiedian and Kuzara 1995). When a new program manager is introduced to an existing contract, he typically begins with no system knowledge and is not immediately able to evaluate the software effort. This results in a lack of corporate memory and continuity. In this case,

confidence in the maturity and capability of the contractor to work effectively without supervision must be established formally in an evaluation. This example reiterates the importance of performing an SCE when lack of oversight is a potential risk.

It is necessary to select an evaluation tool, but this can be a difficult task. It is often difficult to determine which tool to select, specifically because there are so many. Consistency in the type of tool used is of high importance; therefore, it is important that the US DoD approves and uses only a small subset of the available tools in order to make the results meaningful, comparable, and reusable. In practice, it is best to select a few tools that produce results that can be compared with each other (Acton and Aldrich 2000, Humphrey 1988). Much consideration should be given to the goals of the SCE during acquisition. The results obtained from the process must be usable and comparable, especially if there is any intention to reuse results, as will be discussed later.

2.2. Properly Select and Train SCE Team Members

Apply the SEI's guidelines for formation of the SCE team as 'musts' rather than 'shoulds'. Complement the SEI's formal SCE team training with contract-specific education.

The underlying requirement of any SCE is that it be reliable (Bollinger and McGowan 1991, El-Emam and Madhavji 1995). Many studies have been conducted regarding SCE reliability and a common conclusion is that consistency is the most important factor. This would mean that many teams may evaluate the same organization, but they should all come up with the same results. Oftentimes, owing to departmental turnover, it is difficult to keep the same teams together. There are other instances when a team may not harbor enough domain-specific knowledge to properly conduct an SCE (Humphrey 1988). The commonality between any well-conducted SCE is that the SCE team of assessors is capable of doing their job. If an SCE team is not capable of creating a useful evaluation, the SCE will not be successful and will not provide accurate insight into the candidate contractor's capabilities. These facts point to the importance of the SCE team selection (Olson *et al*. 1989). However, this is not always an easy task.

According to the SEI, the SCE team should collectively possess knowledge and experience in the following areas:

- Acquisition policies and procedures
- Project management and planning
- Configuration management
- Software design, development, and methodologies
- Quality assurance
- Systems engineering
- Technical requirements of the contract
- Software testing
- Application domain
- Fluency in CMM
- Expertise in contract-relevant KPAs
- Prior SCE experience.

Knowledge of how to properly conduct an SCE is required of each team member (Olson *et al*. 1989). The SEI provides training to SCE team members to better equip them for their role. This training focuses on team interaction, conducting the SCE, and evaluating the results. It includes such topics as team building, preparing for the SCE, conducting interviews, validating observations, and presenting findings. Training that provides familiarization of the CMM, KPAs, and other software engineering is also recommended by the SEI to team members, in order to better understand the process and the analysis to be performed.

While an SCE team must be properly trained, new team members will be able to gain experience while training. For this reason, it is highly recommended to include SEI-accredited team members who can help train inexperienced teammates. This procedure will instill confidence that the team will properly represent management and CMM expertise (Paulk 1999). However, a team comprising only outside resources, such as SEI-accredited evaluators or members from and external US DoD office, may not represent the interests of the US DoD itself. Some members of the team must be US DoD representatives who are familiar with the specific needs of the US DoD, such as compliance with acquisition policies and end-customer requirements. Without these influences, the candidate contracts may not be properly evaluated in terms of the US DoDs needs (O'Connell and Saiedian 2000). This fact also applies to the topic of the earlier section regarding whether an SCE needs to be performed. A team member representing the interests of the US DoD

would easily recognize these issues, illustrating the importance of proper team member selection.

The US DoD has identified the lack of training in its acquisition and SCE efforts. The Defense Acquisition University (DAU), the Information Resources Management College (IRMC) at the National Defense University (NDU), and the Defense System Management College (DSMC) have collaborated to develop course curriculum to address the US DoD software acquisition management. The courses have been identified by the Under Secretary of Defense for Acquisition and Technology to be integral to the education of acquisition work force personnel. Each course is designed to address a specific acquisition effort. The intended students are those personnel who acquire, develop, engineer, test, evaluate, research, and procure software-dependent systems. Because they address the needs of the DoD, they are highly relevant to SCE team members. The Defense Acquisition Board Science and Technology Committee reported the following:

Improving software education and training is critical – there is a need to coordinate efforts – to integrate software acquisition and development programs into existing courses and to establish mandatory software engineering education for all DoD technical and contractual personnel involved in the acquisition process.

While there is an acquisition career field within the DoD, acquisition offices commonly borrow unrelated personnel and train them for a temporary acquisition assignment. It is not uncommon for a pilot, for example, to cross-train in this drastically different area in order to achieve his own career goals. These individuals are not proficient in acquisition and do not obtain the required depth of acquisition knowledge.

There are also several innate human traits that contribute to a good evaluation team member. For example, consider an assessment team performing an evaluation on a candidate contractor. The team will typically uncover three to five key practices where alternate implementations are used to address the goals of that KPA (Paulk 1999). It takes the intelligence and experience of team members to interpret whether the candidate contractor has implemented his solution in a manner that is adequate. Knowledge and experience are required in order to make an informed judgment about each KPA in its process goals, which are the basic building blocks of the CMM.

2.3. Conduct Parallel Evaluations Performed by the SCE Team and the Contractor

Both the SCE Team and the Contractor should conduct the same evaluation.

There are many opportunities for the potential contractors to put their best foot forward in a way that can deceive most SCE teams. This deception is largely due to the insufficient amount of time given to the SCE team to perform an evaluation (Acton and Aldrich 2000). During the typical week spent on the evaluation, the contractor is asked to provide sample projects and make certain teams available for interviews. The projects they provide may be biased to show more capability that is truly possessed.

When the SCE team performs an evaluation, they have a limited amount of time to perform evaluations of potential contract candidates. During this time, the team must perform several duties (El-Emam and Madhavji 1995):

- Prepare for the evaluation
- Collect data by performing interviews
- Consolidate data
- Make judgments based on the data and interviews
- Report the results and recommend follow-up actions.

A short time frame is often not adequate to perform this amount of evaluation. There is also a significant impact on the understanding of the contractor and the evaluated materials. However, the contractors know this and may exploit the lack of understanding by becoming intentionally vague or overdetailed.

Another shortcoming to the evaluation process is not sampling enough projects. The SCE team may be presented the best of what the contractor has to offer, while the majority of the contractor's projects may not meet the requirements. It is possible to counter this by sampling more projects, but once again scheduling becomes a problem since there is limited time to complete the evaluation. A possible solution would be to have an extra assessor to serve as a floater (Acton and Aldrich). If there are several DoD SCE teams, having an extra person who can move between the teams may help when a team falls behind schedule. The same concept applies to the company performing the internal SCE.

In order to sample more projects and gain an extra layer of insight into the inner workings of

the currently evaluated company, the candidates should provide their own SCE. At first glance, it would appear as though this opportunity would provide the candidate to turn the evaluation into more of a marketing ploy. The candidates could once again make use of intentional vagueness and extreme detail in order to make their past projects look complete. Preventing this possible misuse of the SCE method is done by conducting the evaluations in parallel. The SCE team may choose to evaluate a couple of projects that the candidate evaluates as well, but the SCE team may also focus on other projects that might not be as large (Humphrey 1988). They can then use their evaluation as a baseline and compare their version to the one submitted by the candidate. The results from the comparison may provide evidence of any SCE method abuse or manipulation such as vagueness, nonexistent documents, overdetail, or coaching.

On the same token, performing an accurate SCE internally is beneficial for the candidate contractor, and they are therefore less likely to embellish facts due to the benefits they stand to gain from the demonstrated maturity of an internal SCE. As a case study, Motorola employs self-assessments to motivate and evaluate their in-house processes (El-Emam and Madhavji). In the interest of improving the pool of candidate contractors, it is once again beneficial to require the contractor to perform internal evaluations. The notion of parallel evaluations being performed is also justified by the payoff it may have later during the project. If a potential contractor is unable to perform a self-evaluation before being awarded the contract, it is unlikely that they will be able to do so after they have taken the contract. The parallel evaluation will simultaneously provide an evaluation of the candidate's ability to adhere to later project requirements.

The SCE team may also choose to utilize subtle evidence provided by the candidate contractor's self-evaluation such as key personnel and key projects. This information is provided by the projects presented and the personnel who contributed to the evaluation (El-Emam and Madhavji). It will become obvious to the SCE team which projects the contractor wants to display the most, as well as the personnel who understand their own internal processes (or how to manipulate them) the best. During the SCE team interviews, the employees who did not heavily contribute to the self-evaluation should be sought in order to determine the validity of the contractor's claims.

If the SCE team finds the contract to have provided an evaluation that is similar to the baseline evaluation conducted, then it is more likely that the candidate contractor has been truthful. If the SCE team is comfortable with the self-reviews findings, then the additional projects reviewed may provide the team with much needed information regarding the candidate contractor's procedures. However, if there is a large amount of deviation between the baseline and the self-evaluation, it may be ascertained that the candidate was attempting to take advantage of the opportunity, which may influence the decisions of the SCE team.

Overall, this recommendation may be useful to determine the future performances of contractors on the basis of their abilities to self-evaluate and self-improve. This factor is paramount when contractors will be required to perform many more evaluations during the project, or especially when process maturity is a factor.

2.4. Evaluate a Larger Group of Projects and Interviewees

The SCE team should select a larger group of projects to evaluate. Additionally, the SCE team should have the leverage to select which individual members of those projects will be interviewed.

One of the most important activities performed during an SCE is the evaluation of the candidate contractor's projects. This gives the SCE team a good look at what the contractors have done in the past and what they are capable of providing. It is important that the results of the project evaluations be accurate and useful and that the SCE team gets to see truly representative projects, not just the best of the best.

During a typical SCE, a candidate contractor will submit 10 to 12 project profiles for the team to choose from for evaluation. Once the SCE team has decided on the five or six projects they want to examine, the contractor will then select the personnel who will represent that project. Some guidelines for project selection are as follows (Saiedian and Kuzara 1995):

- The project should represent the software of most concern to the organization.
- The project should represent the software process used in the organization as a whole.

- The project should have staff of at least four members and a life span of at least six months.
- The projects selected should be in different phases of development. *The projects should vary in size and duration.
- The project should not be included if an assessment team member participates in or manages it.

When examining such a small subset of projects, it is easy for the contractor to submit 'token' projects that make the company look good, and then select individuals that are more familiar with the company's processes and will better represent the knowledge that the SCE team is looking for (O'Connel and Saiedian). In order to reduce the possibility of the contractor only putting the 'best foot forward', the SCE team should require the contractor to submit twice as many projects and the team should evaluate twice as many projects. Additionally, it is entirely possible that some projects will not map well to the CMM. By increasing the number of projects that are evaluated, the chances of selecting projects that are difficult to evaluate (and have limited CMM value) will be minimized.

Another potential problem with not evaluating enough projects is the evaluators themselves not seeing everything. As stated earlier, team members must be able to interpret results. Sometimes KPA goals are reached via alternative methods (Paulk 1999). It is important that the evaluation team see enough projects for these methods to become clear. An increase in the number of projects evaluated may also help eliminate the natural human biases. If more evidence is available to provide the evaluator with a complete picture of the process and convincing proof, the possibility for a bias, in any direction, to form is much less likely.

In order to simplify the project profile evaluation process, the submitted profiles should only contain abstract descriptions of each project, such as project size, development language, and contract type. The profiles should also state whether the project has been evaluated before during previous SCEs. By making this a requirement, it may be easier to identify 'Golden Teams' or projects that are perfect for use during an SCE (O'Connell and Saiedian 2000). This knowledge can prompt an SCE team to more closely evaluate the responses and artifacts of the project to ensure that they represent a true maturity, and not one embellished for evaluation purposes. Additionally, the SCE team will be able to evaluate the continuous improvement of the contractor on the basis of how he has incorporated the results of previous SCE efforts into his software process improvement program.

Another way to avoid the contractor submitting 'Golden Teams' during an SCE is for the team to select the project members to interview (O'Connell and Saiedian 2000). The candidate contractor should provide organizational information regarding development staff as well as resumes. This is plenty of information for the SCE team to select an unbiased cross section of staff members. The SCE team should make sure that it selects proper representation of the candidate, which should include expert and novice staff, management, developers, testing personnel, quality assurance, new hires, and senior staff. This procedure will also reduce the effects of 'coached' personnel. A truly mature company should have a staff where each member is fully aware of the company's policies and procedures, and is able to represent the company maturely. By controlling the diversity of the interviewees, the SCE team will have greater ability to achieve a comprehensive, complete, and accurate evaluation.

2.5. Perform Walk-throughs

Complement the documentation review and interviews with a walk-through of the contractor's work environment.

According to SCE 3.0, once the first major steps of interviews and documentation reviews are complete, the SCE team must analyze and consolidate this information. This consolidation process describes the team's progression toward validation of CMM topics by transforming all collected data into a set of strengths and weaknesses, relative to the CMM. This is a time-consuming process that focuses on organizing the information into a manageable sum. During this organization, it is difficult for the team to obtain a complete understanding of the project, and information provided by the contractor may be accepted as true without further verification. If sufficient information is not provided to make a judgment call, the team must decide how to obtain the needed information. The consolidation activity is expected to repeat as many times as necessary until a conclusion can be reached.

Because it is based entirely on documentation review and interviews, the SCE method is described as 'paper intensive'. If the SCE team has not gathered sufficient information to make judgments, it is generally either because no second source has confirmed a practice, use of an alternative practice has not been recognized, or simply that the contractor is not compliant. The SCE team has no mechanism to resolve the discrepancy other than through additional documentation review or reinterviewing personnel.

During a normal SCE that is not for source selection, the evaluation team will present its findings to the candidate contractor in draft form. This gives the contractor an opportunity to disprove any deficiencies that the SCE team has concluded. The draft findings presentation gives the contractor the 'benefit of the doubt' because the SCE team might have misunderstood a concept or overlooked documentation. The contractor can provide additional documentation or personnel to resolve ambiguities, allowing it a final opportunity to present its maturity. This process intends to uncover the true processes that a contractor uses simply by a paper trail and by word of mouth. There are many opportunities for deception on behalf of the contractor, or misinterpretation on behalf of the SCE team (Paulk 1999).

In order to address these problems, a 'walk-through' should become part of the information-gathering process. Any on-site visit allows the SCE team to observe the contractor's processes in action without a change to bias them. These can be conducted spontaneously or as a planned demonstration to the SCE team (O'Connell and Saiedian 2000). For example, if the SCE team has been unable to reach a consensus on the project's configuration management maturity, it might request a demonstration of the configuration control software used to maintain the software baseline. Or, the SCE team might ask to attend a peer review or Software Engineering Process Group (SEPG) meeting. Regardless of application, the walk-through will allow the team to observe and gain a better understanding of the candidate contractor's processes.

Another common scenario that walk-throughs can address is the contractor's use of alternative practices (Paulk 1999). Uncovering these alternative practices can take time, and since thoroughness of the SCE is proportional to the time spent, it is important to be able to identify these practices without too many iterations of the consolidation activity. A walk-through will eliminate this risk because observations can provide an immediate confirmation of contractor compliance. Though a walk-through can aid in uncovering information, it is possible to incorporate them earlier in the SCE method as a complement to the interview and documentation review. This would enable the walk-through to be more encompassing and to interact with more random interviewees selected by the SCE team. Observation of the candidate contractor's teams in action is also possible during a walk-through, and will help the SCE team to get 'proof' of the use and understanding of said processes.

The ability to observe the contractor's processes in action and randomly interact with the employees will help limit the ability of a contractor to embellish its capabilities in the manner described in earlier. This practice will also contribute to the amount of information gathered during the interview process, if used that early in the SCE. Many benefits are possible when using a walk-through as a component of an SCE.

2.6. Reuse Previous SCE Results

Limit evaluation of contractors who have been previously evaluated. All SCE-derived data should be reported to the DoD repository of source selection SCE information for use in future acquisition efforts.

For at least the past decade, industry has criticized the DoD's consistent requirement for an SCE in contract competition. Lt. Col. Vonda, a United States Army acquisition reform staff officer for the Office of the Deputy Under Secretary of Defense affirms, 'Industry continues to assert that it seems as if the government is constantly looking at them, asking the same questions and getting the same answers.' Continual SCEs waste time, money, resources, and inhibit the ability of the contractor to focus on its 'real' job. This common abuse of the SCE process has been termed *redundant reviews*.

In July 1997, the DoD acknowledged its overuse of the SCE. The acting Under Secretary of Defense for Acquisition and Technology directed the System Engineering Steering Group to investigate ways to improve the US DoD's use of SCEs in acquisition. In January 1998, the steering group's efforts resulted in a policy to reuse the results of a previous evaluation whenever feasible.

Two US DoD organizations have been given responsibility to accommodate this policy. The Software Center of the Defense Contract Management Command (DCMC) will be responsible for collecting all information on source selection SCEs conducted on US DoD contractors. The Air Force Electronic Systems Center (ESC) will maintain a repository of US DoD SCE documentation. The information in the repository will be regarded as source-selection sensitive and company proprietary, and will be made available only to government-sponsored evaluation teams.

With this new repository in place, the evaluation process will change. The first task during an evaluation should be to request the most current SCE results for a particular candidate contractor from the ESCs SCE repository. At this point, the acquirer will review all existing data and determine if it matches the requirements of the contract. It should be analyzed for objectivity, timeliness, relevancy, and consistency.

It is important to note that the CMM is an evolutionary model. In order to be promoted a level, the lower levels must be completely satisfied. If a contractor has obtained a certain level in the CMM, it is unreasonable to expect that they will degrade their ranking in future SCEs. If there is sufficient data, another SCE does not have to be performed. If more information is needed, a full SCE can be requested, but this need not be based on the earlier statement that contractors are expected to be on the same maturity level as they were in the previous SCE. The acquirer can then request a 'delta' SCE, which measures only the current capabilities by reviewing the deficiencies identified in past SCEs. As more of these deficiencies are addressed, the contractor can eventually be promoted to a higher maturity level.

The US DoDs SCE repository is still in its infancy. In 1997, a pilot program as ESC tested the concept. It estimated that use of the repository resulted in a $1 million cost avoidance (AR Today). The largest challenges facing ESC and DCMC are education of the US DoD community on the existence and use of the repository in addition to proving its criticality and encouraging its use.

Beyond just the US DoDs SCE repository, the SEI recommends the use of a Process Database to store critical information about used processes. By keeping an extensive collection of data, contractors can keep procedures and metrics in a central location in order to improve overall organization knowledge. Having a repository like this may also assist SCE teams.

Strict, consistent application of the SCE is vital for this reuse initiative to become successful. The data must be accurate and properly derived, as well as the team performing the SCE being qualified and competent. Specifically, the SCE team must represent all relevant areas of expertise and projects and personnel should be carefully selected. This requirement also extends to the thoroughness of any full evaluation. More projects must be sampled and a better understanding must be obtained of a contractor's processes in order to make the reused data useful. This makes it important to conduct walk-throughs, as discussed earlier, in an effort to make sure that the baseline SCE is as complete and accurate as possible, otherwise more resources will be wasted when conducting another full SCE since the previous information was not complete.

2.7. Monitor Performance of Existing Contracts

Continue to evaluate the maturity of the contractor after contract award.

The primary metrics that the SCE derives are capability and maturity of a contractor before awarding the contract. Once the contract has been awarded, it is just as important to monitor current contracts with the SCE (Humphrey 1988). This process is much simpler than any initial SCE since the only contractor project that must be reviewed is the current one.

The US DoD should remain actively involved after the award of the contract to ensure that cost overruns, schedule slips, and poor product quality are avoided. At a minimum, every US DoD contract should include mandatory status reporting and government oversight. The contractor should also be required to maintain and continuously improve its CMM process capabilities, as they relate to the contract. This can be accomplished in two ways. First, the contractor can be required to periodically reassess itself. Second, the US DoD can monitor the contract by performing an additional, contract-specific SCE. Each of these alternatives is described below.

It may be part of the contract that the contractor must periodically reassess himself (O'Connell and Saiedian 2000). Since this reassessment is going to take place during the current contracted work,

the review should focus on that specific contract. Another possibility is that the contractor may be required to address any deficiencies found during the initial SCE, or certain questions on the maturity questionnaire. Either way, the results should be analyzed by the US DoD contract monitor in order to assess any risk that may be growing during the project.

An alternative to reassessments is that a contractor can participate in a secondary evaluation. The SCE method was designed primarily for use as a source selection tool, but is also appropriate for contract monitoring (Besselman *et al*. 1993). Once the contract work has begun, an SCE that focuses on the actual team performing the work can be conducted. This is a slightly different process due to the fact that US DoD representatives will be present during the capability review. This added degree of oversight allows the US DoD to mitigate risks introduced by the contracted team as soon as possible. Use of the secondary SCE is also encouraged by the SEI for long-term relationships of government and contractors.

It should be part of the requirements in the contract that either self-assessments or secondary SCEs be performed during the course of the project. This will hold the contractor more accountable for their claimed capabilities as well as help the US DoD identify and mitigate risks as they evolve. This will also assist the contractor in correction or mitigation of any deficiencies identified during the initial SCE that relate to the contract. As incentive to the contractor and insurance to the US DoD, the results of the contract monitoring SCE or self-assessment and correction of initial SCE-identified problems should be included as considerations for the contract award fee.

3. CONCLUSIONS

The following list summarizes our recommendations for improving SCE practices:

1. *Base the decision to use the SCE on potential risks*: Only perform an evaluation when the contract represents a significant degree of risk. Study the characteristics of the SCE to determine its appropriateness.
2. *Properly select and train SCE team members*: Team members must possess current technical, managerial, and acquisition experience. The CMM knowledge of the SCE team should be complemented by contract-specific expertise.
3. *Conduct parallel evaluations performed by the SCE team and the contractor*: Have the contractors evaluate themselves at the same time as the SCE team evaluations. This will allow the SCE team to synthesize a baseline evaluation and compare results.
4. *Evaluate a larger group of projects and interviewees*: Double the number of projects and personnel that are evaluated.
5. *Perform walk-throughs*: Tour the contractor's work area and conduct spontaneous interviews.
6. *Reuse previous SCE results*: Utilize the US DoD repository of contractor SCE data.
7. *Monitor performance of existing contracts*: Continue to evaluate the CMM initiatives of the contractor after contract award.

Though it is not necessary to implement each one of these recommendations, positive results can come from their combined use. For example, performing walk-throughs alone will help reduce interview and documentation time. However, performing the walk-through and reusing results will gain more time for the overall process and make it possible to focus on specific aspects, such as the walk-through.

Benefits also exist in making a commitment to following these practices. If an organization commits to monitoring performance of existing contracts, a commitment should also be made to reusing previous results. This will form a cyclic process that will not only improve the contractor's internal performance but the U.S. DoD will also benefit by getting a higher quality product.

The most important factor when deciding how to overcome the problems of the SCE is scrutiny of current processes. There may be weaknesses in any organization's SCE process that exposes some of the problems mentioned in this article. If weaknesses can be identified, corrective action should be taken. However, there is still 'no silver bullet' to SCE process problem resolution, but there are options for corrective actions.

ACKNOWLEDGEMENTS

The author wishes to thank Emilie O'Connell for her initial research and contribution to this work. The author also wishes to thank Jake Chambers for his assistance.

REFERENCES

Acton D, Aldrich A. 2000. *Simplifying Assessments*. Lockheed Martin Mission Systems Report.

Besselman J, Byrnes P, Lin C, Paulk M, Puranik R. 1993. *Software Capability Evaluations: Experiences from the Field*, SEI Technical Review'93.

Bollinger T, McGowan C. 1991. A critical look at software capability evaluations. *IEEE Software* **8**(4): 25–41.

El-Emam K, Madhavji N. 1995. The reliability of measuring organizational maturity. *Software Process Improvement and Practice Journal* **1**(1): 3–25.

Humphrey WS. 1988. Characterizing the software process: a maturity framework. *IEEE Software* **5**(3): 73–79.

O'Connell E, Saiedian H. 2000. Can you trust software capability evaluations? *IEEE Computer* **33**(2): 28–35.

Olson T, Humphrey W, Kitson D. 1989. Conducting SEI-assisted software process assessments. Technical Report CMU/SEI-89-TR-7, Software Engineering Institute, Carnegie Mellon University: Pittsburgh, Pennsylvania.

Paulk M. 1999. Using the software CMM with good judgment. *ASQ Software Quality Professional* **1**(3): 19–29.

Saiedian H, Kuzara R. 1995. SEI capability maturity model's impact on contractors. *IEEE Computer* **28**(1): 16–26.