

# A Systematic Survey of CMM Experience and Results

James D. Herbsleb  
Dennis R. Goldenson  
Software Engineering Institute, Carnegie Mellon University  
Pittsburgh, PA 15213 USA  
(412) 268-7389, (412) 268-8506  
<jherbsle, dg>@sei.cmu.edu

## Abstract

*The capability maturity model\* (CMM) for software has become very influential as a basis for software process improvement (SPI). Most of the evidence to date showing the results of these efforts has consisted of case studies. We present a systematic survey of organizations that have undertaken CMM-based SPI to get more representative results. We found evidence that process maturity is in fact associated with better organizational performance, and that software process appraisals are viewed, in retrospect, as extremely valuable and accurate guides for the improvement effort. The path was not always smooth, however, and efforts generally took longer and cost more than expected. A number of factors that distinguished highly successful from unsuccessful efforts are identified. Most of these factors are under management control, suggesting that a number of specific management decisions are likely to have a major impact on the success of the effort.*

**Keywords:** capability maturity model, software process improvement, software process appraisals

## 1: Introduction

The Capability Maturity Model\* (CMM) for software [1, 2] plays a major role as a reference model for software process improvement (SPI) efforts in hundreds of software organizations worldwide. The model is used as a standard for appraising the current state of the organization's software process, as well as a guide for identifying and prioritizing the actions comprising the SPI effort.

---

\* Capability Maturity Model and CMM are service marks of Carnegie Mellon University.

## 1.1: The CMM

The CMM describes five levels of process maturity. Each level has a number of key process areas (KPA) that represent the primary issues the organization needs to address in order to mature its process. At the **initial** level of maturity, software projects rely on the skills, and too often the heroic efforts, of individual engineers. The risk of "runaway" projects is high, and developers are forced to lurch from one emergency to another. The KPAs for the **repeatable** level focus on effective project management, which aims to improve the ability to make and meet reasonable schedule and budget commitments. At the **defined** level, improvement focuses on developing tailorable software processes and other organizational assets, so that organizational learning can span projects. Next is the **managed** level, in which priority is given to monitoring software processes. Finally, in the **optimizing** level, quantitative data are consistently used for continuous process improvement.

## 1.2: Evaluating the CMM

The CMM is not without its critics (e.g., [3]). It is sometimes claimed, for example, that adopting the CMM encourages too much bureaucracy, or that the CMM is incomplete or flawed. This debate is partly concerned with scope, policy issues, and conceptual questions, e.g., whether the model harmonizes appropriately with international standards such as ISO-9000. But the debate also focuses on the supposed consequences of adopting the CMM as the basis for SPI efforts. Will the organization get bogged down in red tape or suffer other damage, or will it benefit and show improved performance? Unlike some other improvement models, the CMM does not explicitly require the results of each change to be measured. It is here that valid and reliable empirical evidence is desperately needed.

There have been several published case studies examining the consequences of CMM-based SPI in a number of organizations, including Hughes Aircraft [4], Raytheon [5], Schlumberger [6], and the Air Logistics Center at Tinker Air Force Base [7]. A larger study of SPI in 13 organizations showed improvements in cycle time, defect density, and productivity [8]. The data also indicate high ratios of benefits to cost, all figures in the range of 4:1 to almost 9:1. As these studies indicate, the evidence is accumulating that CMM-based software process improvement appears to be paying off, at least for some organizations.

The biggest weakness in the evidence to date is *representativeness*. It may be that only successful organizations publish case studies, and will only share data from divisions and projects that have been successful. (See [8], pp. 7-8, 41-47 for a discussion of this and other difficulties in interpreting the data collected to date.) We also need to continue to look beyond defense contractors at commercial and government organizations, to see if their results and problems are comparable. A great deal remains to be learned about how useful CMM-based appraisals actually are, and how successfully software organizations implement their process improvement efforts following their appraisals.

Assuming that some organizations are more successful than others, it is important to understand the enabling and inhibiting factors, particularly those under management control, so that more organizations can become more successful in their process improvement efforts.

### 1.3: This survey.

As part of our ongoing effort to empirically test the critical ideas and claims about the CMM, we designed this survey specifically to address the most significant limitations of previous studies. A carefully conducted survey is an effective way to look at a broad cross-section of a population. In particular, it offers the possibility of examining the results experienced both by highly successful and less successful organizations. The goals of this survey were 1) to find out what typically happens to SPI efforts after assessments, 2) to learn as much as possible about the reasons for success or failure, and 3) to see if the performance reported by more mature organizations is in fact superior to the performance reported by less mature organizations.

## 2: Empirical methodology

### 2.1: The sample

As a first step in a cycle of process improvement, many organizations conduct an appraisal of their software processes, using the CMM as the reference model. The Software Engineering Institute (SEI) maintains a database of appraisal results, which it receives directly from software organizations who have conducted internal appraisals, and from independent vendors. All were performed by authorized appraisers, indicating that they are familiar with the CMM and its associated appraisal methods. The database now has well over 400 appraisals.

We limited the sample to appraisals conducted no less than one year ago (so that there was time for change to take place) and no more than three years ago (so we could find people able to give good accounts of what happened after the appraisals). We also limited the sample to organizations in the United States and Canada, to avoid the expense and difficulty of trying to communicate across great distances and time zones. In order to get a broad and balanced perspective, we decided to try to contact, for each appraisal, a senior technical person and a project manager as well as a member of the software engineering process group (SEPG).

We found 155 appraisals that met our selection criteria in the SEI database. It required an enormous amount of staff time and effort to sift through incomplete records and track down names and addresses through third-party sources. All told, we were able to obtain contact information about 167 individuals representing 61 assessments.

### 2.2: The questionnaire

The questionnaire consisted of 8 pages of mostly closed-ended questions. The topics covered by the questions included

- the value and accuracy of the appraisal,
- current performance of the organization,
- degree of success in addressing the recommendations that were based on the appraisal, and
- background information about the organization and the SPI effort

## 3: Results

### 3.1: Return rate

Of the 167 questionnaires we sent out, we received completed and usable data from 138 of them, for a return rate of 83%. Return rate is extremely important; if the return rate is low (a rule of thumb

often used is that less than 70% is questionable) there is no way to tell if the results are meaningful. It may be that the people who chose to respond are very atypical, e.g., only those who are very pleased or very angry. Our rate of 83% is very good, and was obtained by a vigorous schedule of reminders, phone calls, e-mail, and re-mailing of surveys.

We also succeeded in obtaining responses from individuals in several roles. Of the 138 questionnaires returned, 47 were from a senior member of the technical staff, 47 were from project managers, and 44 were from members of an SEPG. Interestingly, we found no systematic differences among the responses of these three groups.

### 3.2: Maturity profile

Table 1 shows the distribution of self-rated maturity levels.

Maturity Level	Frequency (appraised)	Frequency (self-reported)
1	111	87
2	14	36
3+	9	15
<b>Total</b>	<b>134*</b>	<b>138</b>

**Table 1. Maturity levels. \*We do not have the actual appraised maturity level for 2 of the organizations; four of the respondents were from these organizations.**

We asked each respondent to estimate the current maturity level of the organization. While we had access to the appraised level from our database, these levels were from 1-3 years old at the time the questionnaire was sent out. Since this is enough time for substantial changes in maturity level, we decided to use the self-rating in our further analyses. These self-ratings are, as one would expect, somewhat higher than the appraised levels. These increases are about what we would anticipate in 1-3 years, given what we know about the time it takes to change maturity level [9].

We broke down the responses according to organization type as well as maturity. Respondents from commercial companies make up 36% of the sample, government contractors 37%, government organizations 22%, and other 5%. As one might expect, given the history of the CMM, the government contractors have been using the CMM for process improvement for a longer time and have achieved a higher overall maturity level (see Table 2).

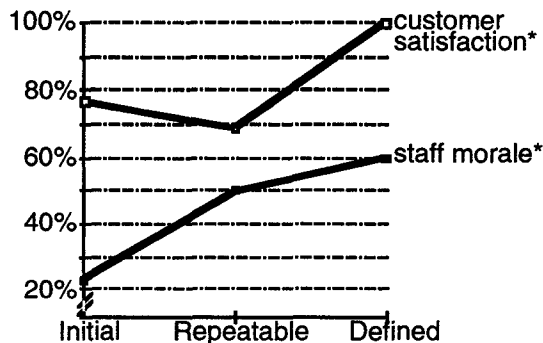
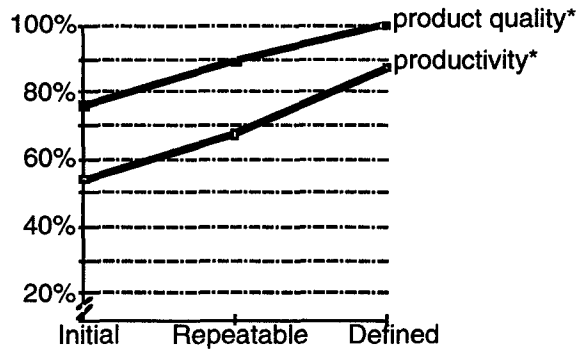
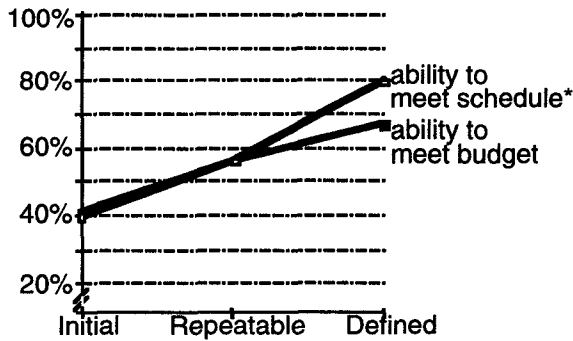
ML	Com-mercial	gov't con-tractor	govern-ment	other
1	40	16	25	6
2	10	20	5	1
3+	0	15	0	0
<b>Tot</b>	<b>50</b>	<b>51</b>	<b>30</b>	<b>7</b>

**Table 2. Maturity Level (ML) by type of organization in the survey sample.**

### 3.3: The impact of SPI

In addition to the maturity ratings, we asked our respondents to indicate where their organization stood on several performance measures, including ability to meet schedule and budget commitments, staff morale, product quality, customer satisfaction, and staff productivity. We do not know the basis on which they made their judgments. All that we need to assume, however, is that if 138 people are asked, e.g., "how good is your ability to meet schedule," that those answering that it is "good" or "excellent" are, on average, better able *in fact* to meet schedules than are those who answer that their ability is "fair" or "poor."

By cross-tabulating maturity level and performance, we can get an indication if performance differs across maturity levels. As Figure 1 shows, those respondents who reported better performance tended to be from higher maturity organizations. One deviation from this pattern is "ability to meet budget," which shows the same pattern as the other dimensions, but is not quite statistically significant. The other deviation is "customer satisfaction," which shows a dip at the repeatable level. The overall pattern, however, clearly shows a relationship between improved performance and higher process maturity.



**Figure 1. Percentage of respondents at each maturity level who rated their performance on each dimension as "good" or "excellent" (the other choices were "fair" and "poor"). The asterisk (\*) indicates that the association between maturity level and performance measure is statistically significant at the .05 level, as determined by a chi-square test.**

Recall from the previous section that the government contractors in our sample have achieved an overall maturity profile higher than the other organizations. In fact, all the level 3 and above organizations in our sample are government contractors. To examine the reported benefits of maturity in the organizations from other sectors, we compared reported performance of organizations at

maturity levels 1 and 2, *excluding* all the government contractors from the analysis. We found statistically significant relations between higher maturity level and ability to meet schedule, ability to meet budget, and higher staff morale. The relationships between maturity level and product quality, customer satisfaction, and productivity are not quite statistically significant. However, *all* of the relationships remain in the same direction as seen in Figure 1. Such consistency is highly unlikely by chance alone.

### 3.4: The appraisals

We asked several questions about the appraisals and the findings and recommendations that were based on them. Two of them asked respondents to evaluate the accuracy of the appraisals in the light of their experience in the 1-3 years that had elapsed since the time of the appraisal. Table 3 summarizes answers to a question about how accurately the appraisal identified the organization's major problems, while Table 4 shows similar data on accuracy of identifying strengths.

Response	% of answers
very accurate	63
generally accurate	35
not very accurate	2

**Table 3. Accuracy identifying problems.**

Response	% of answers
very accurate	38
generally accurate	54
not very accurate	8

**Table 4. Accuracy identifying strengths.**

In addition to these questions about accuracy, we asked respondents directly for their overall judgment of the value of the appraisal. Seventy-four percent agreed or strongly agreed with the following statement:

The appraisal was well worth the money and effort we spent; it had a major positive effect on the organization.

Not only did the respondents report that the appraisal was valuable, in response to another question, 90% reported that much or all of the process improvement effort was in fact based on the appraisal results. In addition, the vast majority of respondents (86%) agreed or strongly agreed that the CMM provided a valuable roadmap to help in establishing the order in which improvements should be made.

Another set of questions asked about a number of criticisms one sometimes hears about appraisals based on the CMM. Table 5 displays the percentage of respondents who agreed or strongly agreed with statements about these criticisms.

Statement	% agree
recommendations are too ambitious	29
results too dependent on team	18
not enough room for team to exercise judgment	4
wrong projects/participants	8
people not fully honest	6

**Table 5. Potential problems with appraisals.**

All in all, these responses are quite positive. In the light of 1-3 years of subsequent experience, the vast majority of respondents still believed the appraisal results to be accurate, useful, and worth the expense. A substantial minority of respondents did say the recommendations coming out the appraisal process were too ambitious and about twenty percent expressed concern about results being too dependent on the judgment of the appraisal team. But these concerns are not widely shared, and fewer than 1 in 10 of the respondents agreed with any of the other critical statements about appraisals.

### 3.5: Progress since the appraisal

We asked a number of questions to try to determine how successful organizations have been, after the appraisal, in actually taking effective action to improve their software processes. Table 6 shows the distribution of responses to a question asking how successful they have been in addressing the recommendations based on the appraisal. Clearly, the degree of perceived success varies considerably.

Response	% of answers
little if any thus far	14
limited success	30
moderate success	26
substantial success	23
marked success throughout the organization	8

**Table 6. Percent of respondents reporting various degrees of success. (Total > 100 because of rounding error.)**

Another important measure of progress is just how far in the typical sequence of post-appraisal activities the organization has gone. After the appraisal, an organization typically puts together an action plan, forms teams to carry out the plan, implements changes in pilot projects, then rolls out changes throughout the organization. As we can see in Table 7, nearly all the organizations completed an action plan and established teams to carry it out. About half implemented process changes on a pilot basis, and about half rolled changes out through the organization.

Response	% of yes answers
create action plan?	96
establish action teams?	89
changes in pilot projects?	53
changes throughout organization?	53

**Table 7. Percent of "yes" answers to questions about progress of SPI.**

If we break down the last two rows in Table 7 in a slightly different way, we see (Table 8) that about a third (35%) of organizations implemented changes both in pilots and throughout the organization, while 18% percent had only implemented changes in pilots. Over a fourth (29%) of the organizations had not implemented any changes yet. Another 18% implemented changes throughout the organization, apparently without conducting pilots.

Changes in Pilot	Changes throughout	
	yes	no
yes	35%	18%
no	18%	29%

**Table 8. Cross-tabulation for answers to questions about change in pilots and change throughout organization.**

Although there are substantial differences among organizations in the amount of progress made, very few (less than 4%) of the respondents said that the SPI efforts were counterproductive, i.e., that things had actually worsened. And 90% disagreed or strongly disagreed that focusing on the CMM had caused the organization to *neglect* important issues. About a quarter (26%), however, agreed with a statement that "nothing much has changed," and just short of half (49%) said there had been a lot of disillusionment over the lack of results.

There is little evidence that these frustrations are due to the most loudly-voiced criticisms, however. For example, 84% disagreed or strongly disagreed that software processes had become more rigid or

bureaucratic. In the commercial and government sectors, respondents from higher process maturity organizations reported that it took *less* "paperwork to get things approved" than respondents from lower maturity organizations. (The 2 variables were unrelated for government contractors, where paperwork requirements are largely driven by contractual requirements.)

More frequent were difficulties related to time, effort, and resources. Not surprisingly, a number of organizations (42%) said that SPI had been overcome by events and crises, with other things taking priority. And a whopping 72% said SPI had suffered due to time and resource limitations. It also appears that a number of organizations had fairly unrealistic expectations – 77% said that SPI took longer than expected, and nearly as many, 68%, said that it cost more than expected. Recent data indicate that it generally takes about two years to move up a maturity level [9].

The other major source of difficulty identified in the survey is the need for more direction about how to actually *do* the process improvement. In one question, two-thirds (67%) of the respondents agreed or strongly agreed with the statement, "We understood what needed to be improved, but we needed more guidance about how to improve." And more than half (57%) agreed or strongly agreed that they needed more individualized mentoring and assistance.

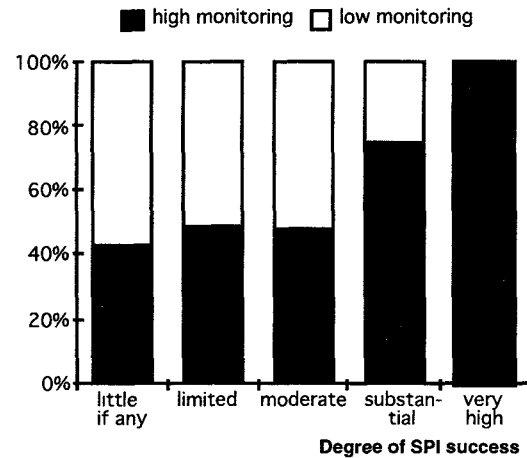
### 3.6: Who succeeds?

In addition to questions about SPI and performance, we asked a number of other questions about the organization and about the SPI effort to try to understand why some organizations were very successful and others were not. In this section, we will present characteristics of highly successful and of less successful organizations. Unless otherwise specified, all the relationships below are statistically significant as determined by a chi-square statistic ( $p \leq .05$ ).

Where SPI efforts tended to be more successful, our respondents tended to agree or strongly agree with the following:

- managers actively monitor progress
- clear, well-understood SPI goals
- clear, compensated assignment of responsibility for SPI
- involvement of technical staff in SPI
- process improvement people highly respected
- staff time and resources dedicated to SPI are good or excellent.

Figure 2 shows an example typical of these relationships. (For more detailed results, see [10])

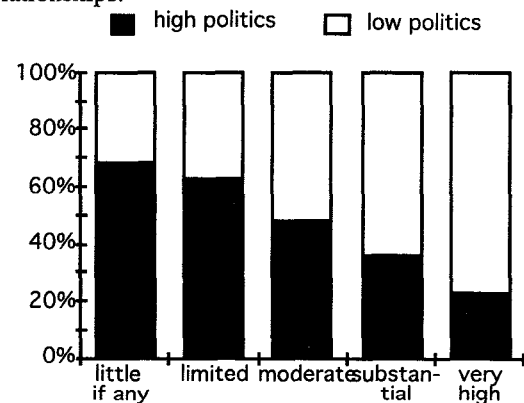


**Figure 2.** For each category of success with the SPI effort, the bars represent the percent of respondents who agreed ("high") or disagreed ("low") that senior management actively monitors SPI progress.

Respondents from organizations with less successful SPI efforts tended to agree or strongly agree with the following:

- lots of organizational politics
- lots of turf guarding
- discouragement and cynicism from previous unsuccessful improvement experience
- belief that SPI gets in way of real work
- need more guidance on how, not just what, to improve

Figure 3 shows an example typical of these relationships.



**Figure 3.** For each category of success with the SPI effort, the bars represent the percent of respondents who agreed ("high") or disagreed ("low") that their organizations are characterized by a high level of "organizational politics".

## 4: Management Implications.

Many of the factors associated with success and failure are under management control. It appears that it is very important to invest adequate resources, to make goals very clear, and to actively monitor progress toward those goals. Interestingly, based on responses to other questions, it does not seem to matter whether the manager is perceived as actually understanding the issues faced by practitioners, or whether rewards and incentives for successful SPI are established.

Setting up the SEPG correctly is also critical. It must be staffed with highly respected people who have clear responsibility for SPI, and who engage in SPI as a funded (not spare time) activity. They must be sure to involve the technical staff in the effort.

Some of the factors that appear to block success are characteristics of the organizational culture, and thus are probably very difficult to overcome. For example, turf guarding and organizational politics make it very difficult to bring about change, presumably because change triggers latent conflicts and brings hidden agendas into play. Still, good managers deal with such issues every day, and address them with at least some success.

Several of the results point to the importance of ensuring that SPI is viewed as a high priority. If management fails to convey this message, and SPI is viewed as something that "gets in the way of real work," success is less likely. And since discouragement and cynicism from previous failed efforts is a good predictor of another failure, every effort should be made to "do it right the first time." It may be necessary to get outside help in the form of a consultant or to hire personnel experienced in SPI. In any event, if the responsible people are saying they need more help understanding how to make the improvements, management should listen. The indications are that an unmet need is associated with lower success.

As we saw in a previous section ("3.5: Progress since the appraisal"), setting realistic expectations is important. There is weak support for a finding that overly-ambitious recommendations hurt chances for success. In addition, most respondents underestimated the time and expense of SPI, which resulted in a fair amount of discouragement. The results of SPI overall, we are finding, are quite positive but they are by no means easy to achieve.

There are several other factors about which a cautious manager should be aware, although the evidence for their importance is not strong. An adverse commercial environment (decreasing demand) or turbulent internal environment (reorganization, senior management turnover) may make the SPI effort more difficult. Middle management and

technical staff turnover, however, do not seem to be related to the degree of success.

## 5: Discussion

In this section, we discuss the goals of the study and the extent to which we were able to meet them. We go on to identify important remaining issues.

### 5.1: Goal 1

*Find out what typically happens to SPI efforts after the assessment.*

Compared to previous work, this study is more representative of the software community. It is drawn from both successes and failures, it includes a substantial number of commercial and government organizations as well as government contractors, and it spans several maturity levels. Respondents included approximately one-third senior technical staff, one-third project managers, and one-third SEPG members, so we are not getting just the input of those individuals personally invested in the CMM. As we mentioned above, responses from these three groups did not differ in any systematic way.

We sent questionnaires to every person we could identify who was in a good position to see what happened in the 1-3 years following an SEI-authorized appraisal. We followed up aggressively to get a good (83%) rate of return and avoided self-selection problems.

### 5.2: Goal 2

*Learn as much as possible about the reasons for success or failure.*

We found a number of attributes of the SPI effort, the organization, and the organizational culture that were strongly associated with the degree of success. Many of these are under management control.

### 5.3: Goal 3

*See if the performance reported by more mature organizations is superior to the performance reported by less mature organizations.*

We found a number of important reported differences in performance between more mature and less mature organizations. Unlike previous case studies [e.g., 5, 6, 7, 8] the survey reported here is forced to rely on self-reported levels of performance. This was a deliberate design tradeoff in order to allow us to look at a broader cross-section of organizations. The results we obtained with this method are largely consistent with the case studies in the literature and with our previous work. This exemplifies our tactic of approaching the important research questions with

a variety of techniques that make different assumptions, gradually converging on a set of reliable results.

#### 5.4: Remaining issues

No single study can hope to test all the important ideas and claims about the CMM and what happens as organizations implement SPI efforts based on the CMM. The CMM is a complex reference model, and the effects of organizational change are difficult to measure and difficult to trace back to particular causes.

We are designing a coordinated series of studies to address these issues. The conceptual backbone of these studies is a list of critical assertions about the CMM, i.e., the ideas and claims about the content of the CMM and the consequences of adopting the CMM that most concern the software engineering community. The current draft list includes 15 assertions about

- predictability and performance by maturity level
- order and content of the key process areas
- moving up the maturity scale -- time, cost, and reporting progress
- useability and implementation issues
- factors that influence success

We intend to address all these critical assertions in more detail over the next several years.

One line of work we have recently undertaken is to establish collaborative relationships with a small number of software-dependent organizations in order to help establish consistent measures and track the results of investments in process and technology. The approach is within the spirit of the Software Engineering Laboratory pioneered by Basili and associates [11], but focused primarily on SPI and on determining business results and how to get them. Another future thrust is to attempt to provide more actionable guidance for SPI, focusing in more depth on factors that affect success of improvement efforts within particular key process areas.

## 6: References

- [1] M. C. Paulk, B. Curtis, M. B. Chrissis, and C. V. Weber, "Capability Maturity Model for Software, Version 1.1," CMU/SEI-93-TR-24, Software Engineering Institute, Pittsburgh, PA Feb., 1993.
- [2] M. C. Paulk, C. V. Weber, S. M. Garcia, M. B. Chrissis, and M. Bush, "Key Practices Of The Capability Maturity Model, Version 1.1," CMU/SEI-93-TR-25, Software Engineering Institute Pittsburgh, PA Feb., 1993..
- [3] C. Jones, "Gaps in SEI Programs," Software Development, vol. 3, No. 3, March, pp. 41-48, 1995.
- [4] Watts S. Humphrey, Terry R. Snyder, and Ronald R. Willis, "Software Process Improvement at

Hughes Aircraft," IEEE Software, Vol. 8, No. 4, July 1991, pp. 11-23.

- [5] R. Dion, "Process improvement and the corporate balance sheet," IEEE Software, vol. 10, pp. 28-35, 1993.

- [6] H. Wohlwend and S. Rosenbaum, "Software improvements in an international company," Proceedings of 15th International Conference on Software Engineering, pp. 212-220, Los Alamitos, CA, 1993.

- [7] W. H. Lipke and K. L. Butler, "Software process improvement: A success story," Crosstalk, November, pp. 29-39, 1993.

- [8] J. Herbsleb, A. Carleton, J. Rozum, J. Siegel, and D. Zubrow, "Benefits Of CMM-Based Software Process Improvement: Initial Results," Software Engineering Institute, CMU/SEI-94-TR-13, August 1994.

- [9] W. Hayes and D. Zubrow, "Moving on up: Data and experience doing CMM-based process improvement," Software Engineering Institute, CMU/SEI-95-TR-08, August 1995.

- [10] D. Goldenson and J. Herbsleb, "After the appraisal: A systematic survey of process improvement, its benefits, and factors that influence success," Software Engineering Institute, CMU/SEI-95-TR-09, August 1995.

- [11] V. R. Basili, G. Caldiera, F. McGarry, R. Pajarski, G. Page, & S. Waligora. "The Software Engineering Laboratory - an operational software experience factory." Proceedings of 14th International Conference on Software Engineering, pp. 370-81, Melbourne, Vic., Australia, 1992.

## Acknowledgments

We thank all those who took the time and effort to complete the survey. Without the support and cooperation of these individuals, this work would not have been possible.

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