

A Study of Systems Engineering Effectiveness: Building a Business Case for SE

Joseph P. Elm
Software Engineering Institute, Carnegie Mellon University
jelm@sei.cmu.edu

Copyright © 2011 by Carnegie Mellon University. Published and used by INCOSE with permission.

Abstract. Understanding the value of Systems Engineering (SE) is necessary to justify a project's investment in SE resources and activities. To identify the value of SE, the Systems Engineering Effectiveness Committee (SEEC) of the National Defense Industrial Association (NDIA), in collaboration with the Software Engineering Institute (SEI), developed and executed a survey of defense industrial contractors. The survey analyzed the relationships between the SE activities applied to individual projects, and performance of those projects. Results of this survey, published in 2007 and 2008^[1], indicated strong relationships between many SE efforts applied early in the project and the overall success of the project. Today, NDIA, in collaboration with the Institute of Electrical and Electronics Engineers Aerospace and Electronic Systems Society (IEEE-AESS) and SEI, is engaged in extending that study to build a stronger business case for the application of effective SE to product development projects.

Background

An understanding of the value of Systems Engineering (SE) is necessary to justify a project's investment in SE resources and activities. And yet, quantitative data showing the contributions of SE to project or program success is sparse. Prior research by Gruehl^[2] showed that NASA projects that spent a larger percentage of their efforts in the early phases of the project exhibited smaller cost overruns. Likewise, research by Honour^[3] has shown that projects that devote 15% to 20% of their development budgets to SE exhibit smaller cost and schedule variances. While producing valuable insights into the role of SE in development projects, these studies have been based on a small number of data points, and were not constructed to show the contribution of specific SE activities to project performance. Subsequently, in 2006, the National Defense Industrial Association Systems Engineering Division (NDIA-SED) embarked on a project to collect quantitative evidence of the value of SE.

NDIA SE Effectiveness Study

The NDIA-SED established the Systems Engineering Effectiveness Committee (SEEC) to obtain quantitative evidence of the benefit of SE best practices upon project performance. The SEEC chose to attack this task by developing and executing a survey of individual development projects executed by defense industrial contractors. The hypothesis of the survey was:

The effective performance of SE best practices on a development program yields quantifiable improvements in the program execution (e.g., improved cost performance, schedule performance, technical performance).

The survey identified the SE practices utilized on these projects, collected project performance data on these projects, and statistically measured relationships between the application of these SE practices and project performance. Those practices with the strongest relationship to better project performance could be considered "best practices."

Assessment of SE Effectiveness demands knowledge of both the SE activities performed, and the results produced by the application of these activities. Measurement of both of these parameters is problematic.

Measurement of the SE activities performed is hindered by the absence of a widely accepted definition of what comprises SE; thus, activities that one project considers as SE, another may not. Furthermore, in many organizations, SE is not an independently planned and tracked activity. Instead, it is distributed among and embedded within many other project tasks.

Reliable means of assessing project performance are also lacking. While many projects employ the Earned Value Management System (EVMS), the manner in which it is employed is not consistent. EVMS is calculated from variances from a baseline; therefore, it is highly sensitive to revisions in those baselines. Within some projects, baselines are only revised upon contract change orders. In others, baselines may be changed during replanning activities. In yet others, baselines are revised at fixed intervals. These different baselining methods can produce significant variations in the meaning of EVMS data. Furthermore, EVMS assesses only satisfaction of budgetary and schedule needs. It includes no means of assessing satisfaction of performance requirements.

It was this lack of reliable data that led the SEEC to focus on this survey as a means of examining the effectiveness of SE.

SURVEY DEVELOPMENT AND DEPLOYMENT

The SEEC chose to measure the application of SE activities by assessing the SE-related work products produced by the each project in the survey sample. Due to the absence of a widely accepted definition of what constitutes SE, the SEEC used the Capability Maturity Model Integration[®] (CMMI[®])^[4] as a guide to identify appropriate work products. The SEEC assembled a panel of SE experts drawn from government, industry, and academia. Starting with the CMMI Model v1.1, the panel identified the work products that, in the judgment of these SE experts, result from SE tasks. The presence and the quality of these work products provide an indication of the magnitude of the SE activities performed on the project.

The process of identifying SE work products was as follows:

1. Extract all listed work products from the CMMI.
2. Identify the subset of these work products that (in the judgment of the SEEC SE experts) result from SE activities.
3. Extract those work products that (in the judgment of the SEEC SE experts) are most significant.

The result of this process was a survey questionnaire that included 88 questions addressing 71 SE-related work products. These questions reflected SE activities performed within the following 12 categories:

- Project Planning
- Requirements Development and Management
- Trade Studies
- Technical Solution Processes
- Verification
- Configuration Management
- Project Monitoring and Control
- Product Architecture
- Integrated Project Teams
- Product Integration
- Validation
- Risk Management

[®] “Capability Maturity Model Integration” and “CMMI” are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

Project performance was measured in three dimensions: budget satisfaction, schedule satisfaction, and requirement satisfaction. Budget and schedule satisfaction were assessed based upon reported EVMS data as well other reported variations in project schedule and project cost. Satisfaction of customer requirements was assessed directly. The result of this process was an additional 13 questions addressing overall project performance.

Additionally, the survey examined the degree of challenge presented by the project. Project Challenge (PC) was measured via 44 questions on the survey questionnaire, and considered factors such as:

- the life-cycle span of the project
- sources of technical challenge
- project value and financial stability
- project duration and schedule stability
- the size of the project staffing
- the organizational complexity of the project (e.g., number of organizations involved)
- the completeness and stability of the project requirements
- the stability of the project contract.

The survey questionnaire also included questions probing the environment in which the project was being executed. Factors considered included:

- the acquiring organization
- the end-user for the delivered product
- the deployment environment of the end-product
- the contract type
- the executing organization's process improvement efforts
- the executing organization's process maturity
- the executing organization's prior related experience

The survey was executed via the web, with responses received from August 2006 through November 2006. All responses were anonymous to protect proprietary project data—a necessity to obtain sufficient participation. As an additional protection, data was accessible only to designated staff of the SEI, a non-competitive FFRDC acting as an honest broker of information.

Sixty-four valid survey responses were received; 48 complete responses and 16 partial but useable responses. While the sample sizes are small, they are adequate. Most importantly, they exhibit sufficient variation. Without this variation, the identification of statistical relationships between variables would not be possible.

DATA ANALYSIS

Response data from each project was analyzed to compute a Systems Engineering Capability (SEC) score for each of the 12 aforementioned SE activity categories (i.e., Project Planning, Project Monitoring and Control, etc.). An overall SEC score covering all of the categories was also computed. All scores ranged numerically from 1 to 4, (1 = low capability, 4 = high capability). Likewise, a Project Performance score (Perf) (1 = poor, 4 = good) and a Project Challenge score (PC) (1 = easier, 2 = more difficult) were computed. All of the scores show sufficient variation to enable meaningful analysis of the relationships between them. Distribution histograms for PC, Overall SEC, and Perf are shown in Figs. 1 through 3.

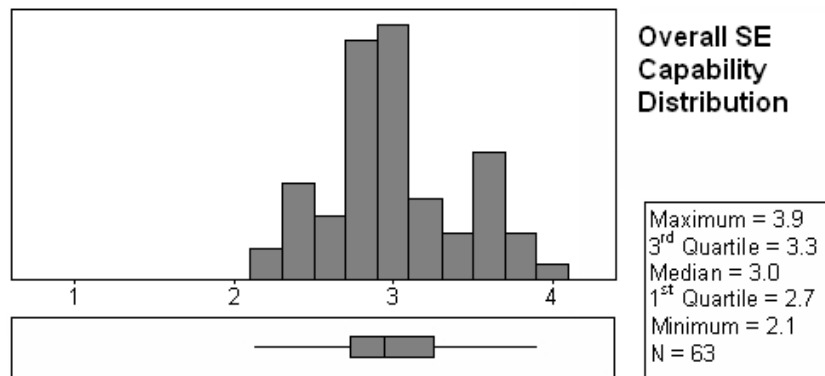


Fig. 1. Overall SEC Distribution

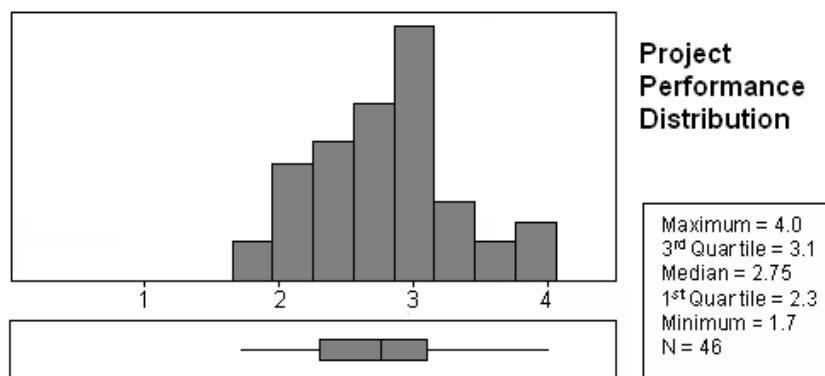


Fig. 2. Project Performance (Perf) Distribution

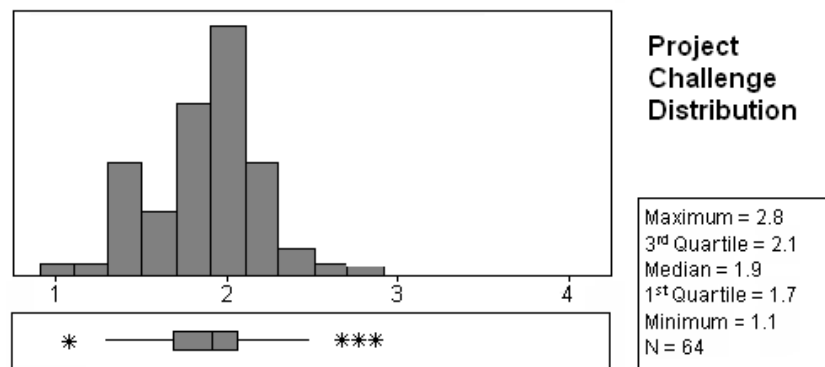


Fig. 3. Project Challenge (PC) Distribution

This data was then analyzed to identify relationships between these scores. For each of the SEC categories, the responses were trichotomized into lower, middle, and upper thirds. Likewise, the Project performance scores were also trichotomized. This enabled the creation of 9-panel mosaic charts clearly showing the distribution of project performance for each of the three subsets of SEC capability. Fig. 4 shows an annotated example of such a mosaic chart.

Goodman and Kruskal’s “gamma”, as shown in Fig. 4, is a proportional reduction in error (PRE) statistic that provides a measure of association that expresses the strength of relationship between two ordinal variables. It varies from +1 to -1, with

- values near -1 indicating a strong opposing relationship
- values near 0 indicating a weak or absent relationship (statistical independence)
- values near +1 indicating a strong supporting relationship

“p,” also shown in Fig. 4, is generally interpreted as the probability that one would observe a statistical relationship in a sample of data by chance alone.

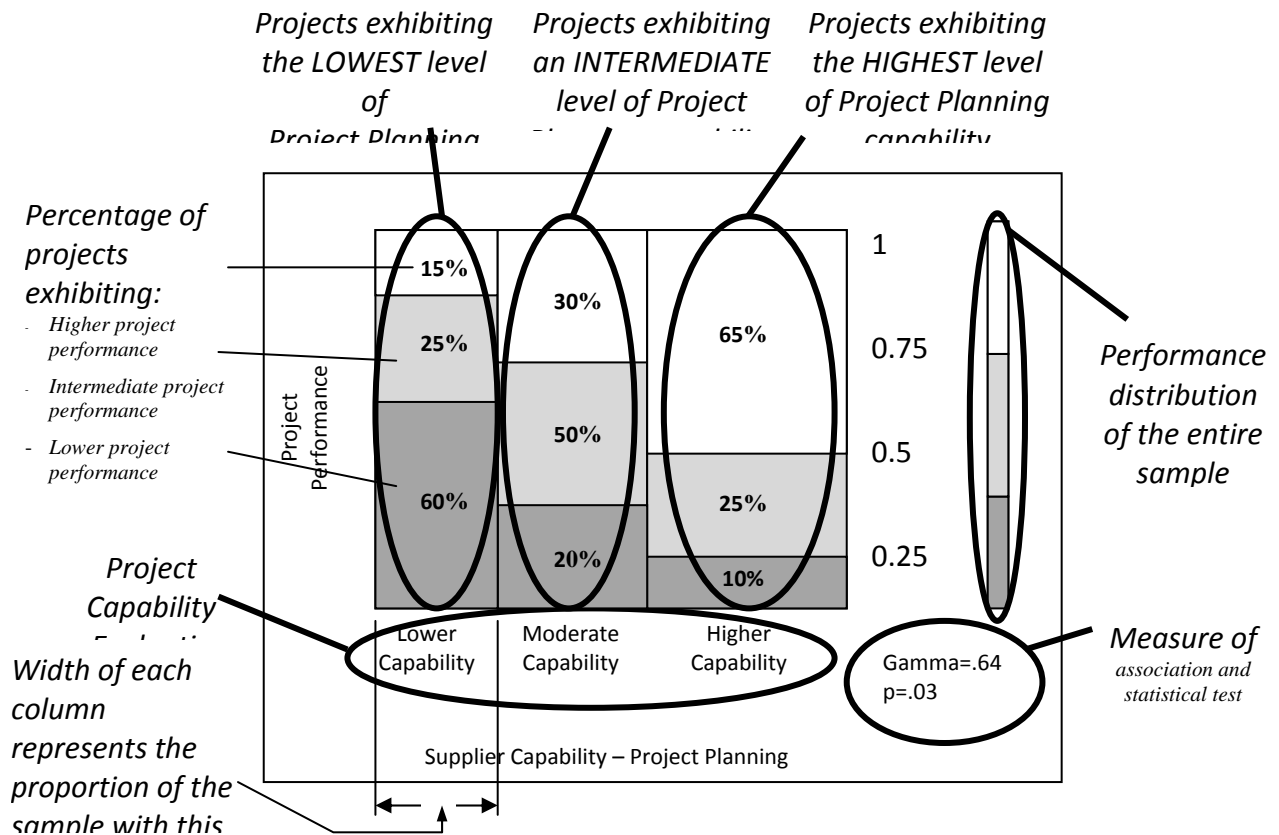


Fig. 4. Example Mosaic Chart

SURVEY RESULTS

The analysis described above was performed for each of the 12 aforementioned SE activity categories, as well as for ensembles of these categories, including an overall measure of SE capability encompassing all categories. This enabled us to identify the relationship of each of these categories to project performance.

Overall SE Capability vs. Project Performance

Fig. 5 shows the relationship between the overall SE capability of the project (Overall SEC) and Project Performance (Perf). Note that for the projects with the lowest overall SE capability, only 15 percent of the projects exhibited the best project performance. However, for projects with the highest overall SE capability, 56 percent of the projects exhibited the best project performance. The gamma value of 0.32 reflects a moderately strong relationship between total SE capability and project performance.

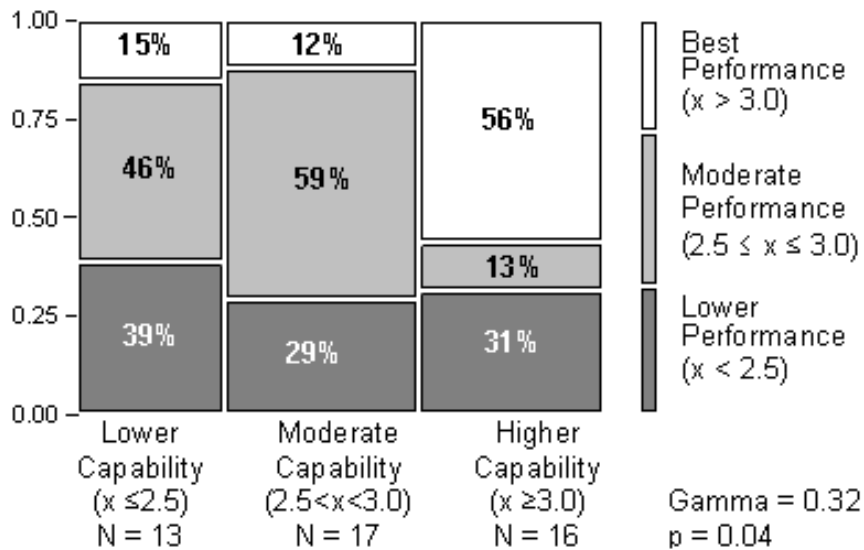


Fig. 5. Relationship Between Overall SE Capability and Project Performance (Perf)

Project Challenge vs. Project Performance

Fig. 6 shows the relationship between Project Performance (Perf) and Project Challenge (PC). As expected, the graph shows that more challenging projects exhibit poorer performance than less challenging projects. For the least challenging projects sampled, 50 percent exhibited the highest project performance. However, for the most challenging projects sampled, the percentage of projects delivering the highest performance dropped to 25 percent. It should be noted that the sample of projects analyzed did not include extremely challenging projects. As seen in Fig. 3, on a scale of 1 to 4 for PC, the median for the represented projects was 1.9, and the maximum was 2.8. Thus, extremely challenging projects are not represented in this sample.

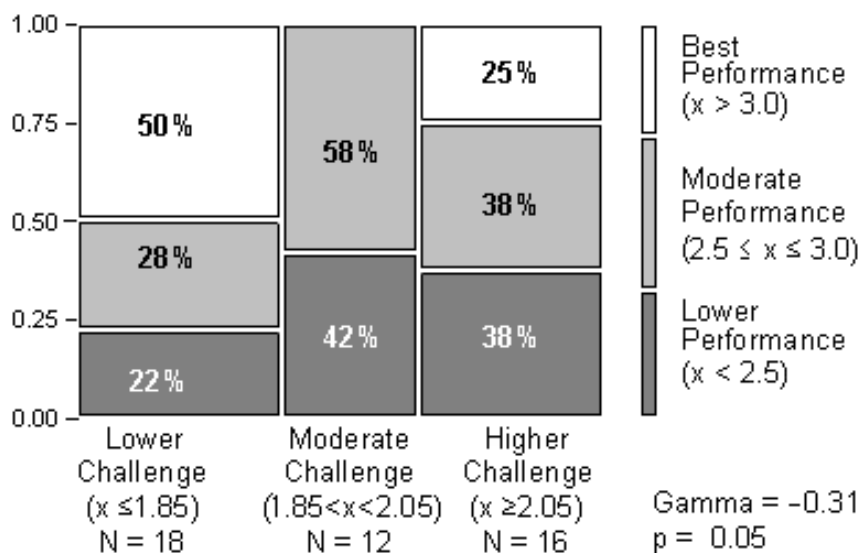


Fig. 6. Relationship Between Project Challenge (PC) and Project Performance (Perf)

One would expect the success rate for these more challenging projects to continue to decrease below the 25 percent shown in Fig. 6.

Product Architecture Practices vs. Project Performance

The impact of the application of Product Architecture practices on Project Performance was analyzed. Assessment of Product Architecture was based upon a series of questions probing

- the documentation, timeliness, management, and accessibility of a high-level product structure using multiple views
- the accuracy, timeliness, management, and accessibility of interfaces descriptions
- the documentation and usage of guidelines for COTS component utilization

As seen in Fig. 7, these Product Architecture factors had a moderately strong relationship (Gamma = 0.40) to Project Performance.

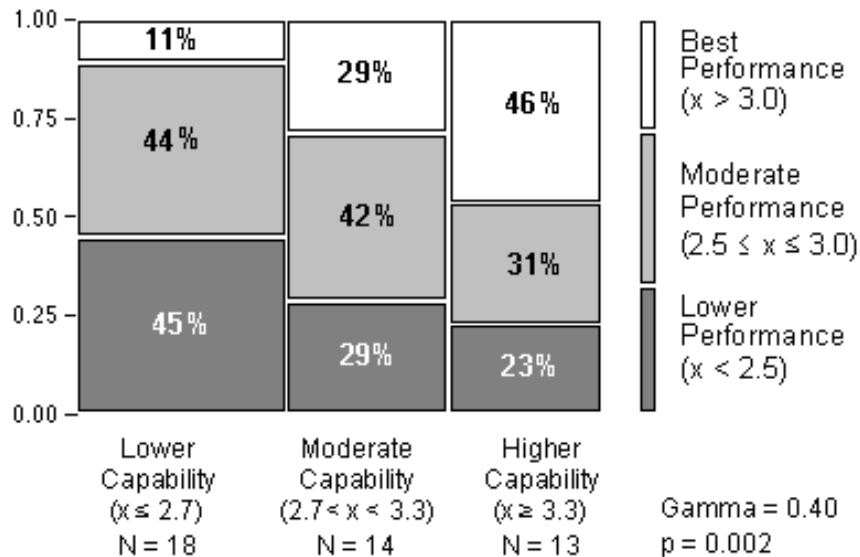


Fig. 7. Relationship Between Product Architecture Practices and Project Performance

For the projects with the lowest Product Architecture capability, only 11 percent exhibited the best project performance. However, for projects with the highest Product Architecture capability, 46 percent exhibited the best project performance.

Trade Study Performance vs. Project Performance

The impact of Trade Study utilization on Project Performance was analyzed. Assessment of Trade Study utilization was based upon a series of questions probing:

- performance and documentation of trade studies between alternate solutions
- utilization of definitive and documented selection criteria
- stakeholder involvement in the development and performance of trade studies
- documentation and accessibility of trade study results

As seen in Fig. 8, these Trade Study factors had a moderately strong relationship (Gamma = 0.37) to Project Performance.

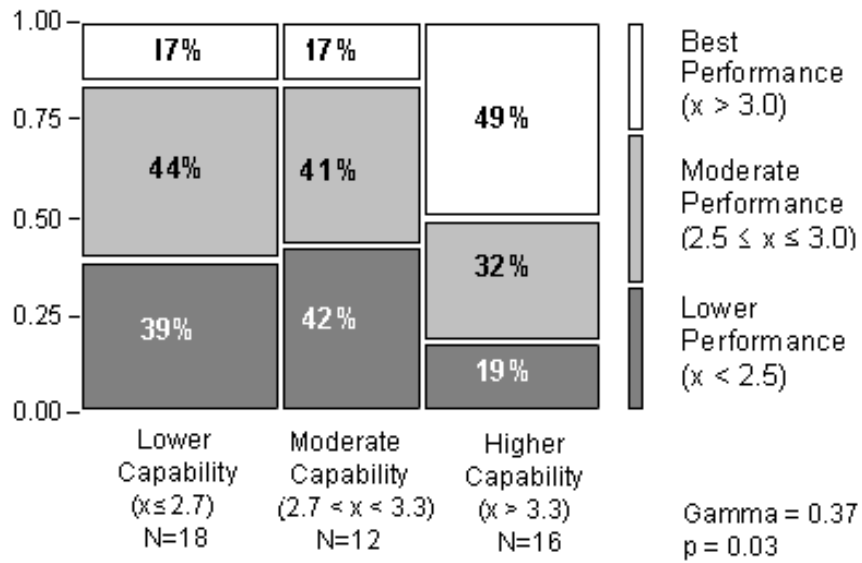


Fig. 8. Relationship Between Trade Study Performance and Project Performance

For the projects with the lowest Trade Study utilization, only 17 percent exhibited the best project performance. However, for projects with the highest Trade Study capability, 49 percent exhibited the best project performance.

Integrated Project Team Utilization vs. Project Performance

The impact of Integrated Project Team (IPT) utilization on Project Performance was analyzed. Assessment of IPT utilization was based upon a series of questions probing

- the use of IPTs
- the effectiveness of IPTs
- supplier participation in IPTs
- inclusion of an IPT with assigned SE responsibility
- SE representation on other IPTs

As seen in Fig. 9, these IPT utilization factors had a moderately strong relationship (Gamma = 0.34) to Project Performance.

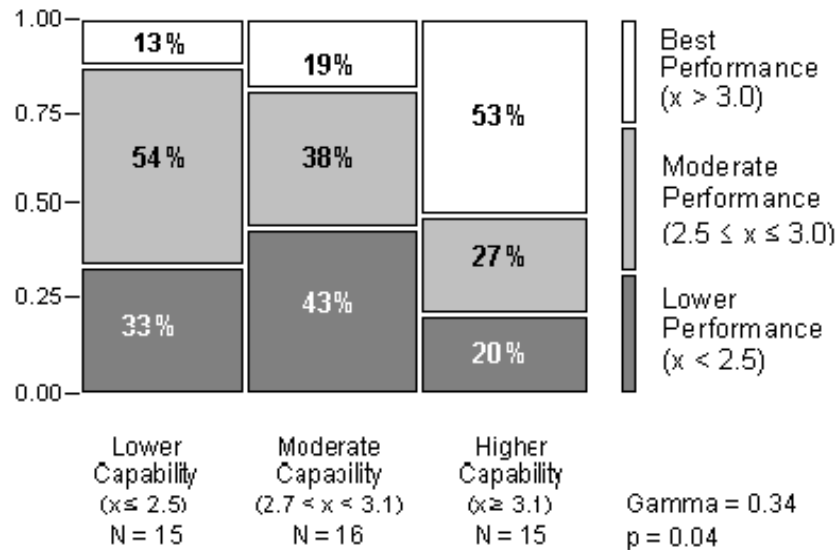


Fig. 9. Relationship Between Integrated Project Team (IPT) Utilization and Project Performance (Perf)

For the projects with the lowest IPT capability, only 13 percent exhibited the best project performance. However, for projects with the highest IPT capability, 53 percent exhibited the best project performance.

Requirements Development and Management vs. Project Performance

The impact of Requirements Development and Management on Project Performance was analyzed. Assessment of Requirements-related capability was based upon a series of questions probing

- the development and documentation of project requirements based on stakeholder needs, expectations, and constraints
- the documentation, timeliness and accuracy of all customer-specified requirements
- the documentation, timeliness and accuracy of all derived requirements
- the documentation, timeliness and accuracy of the hierarchical allocation of both customer and derived requirements to system elements in the configuration baselines
- the documentation, timeliness and accuracy of operational concepts and their associated scenarios
- the documentation, timeliness and accuracy of use cases (or their equivalent)
- the documentation, timeliness and accuracy of product installation, maintenance and support concepts
- the definition and application of criteria for identifying authorized requirements providers
- the definition and application of criteria for evaluation and acceptance of requirements
- the definition and application of a formal process for requirements approval by relevant stakeholders
- performance and documentation of requirements impact assessments for proposed requirements changes
- the usage, timeliness and accuracy of a requirements tracking system
- the management of requirements documents under a configuration control process
- the accessibility of requirements documents to relevant project staff

As seen in Fig. 10, these Requirements-related factors had a moderately strong relationship (Gamma = 0.33) to Project Performance.

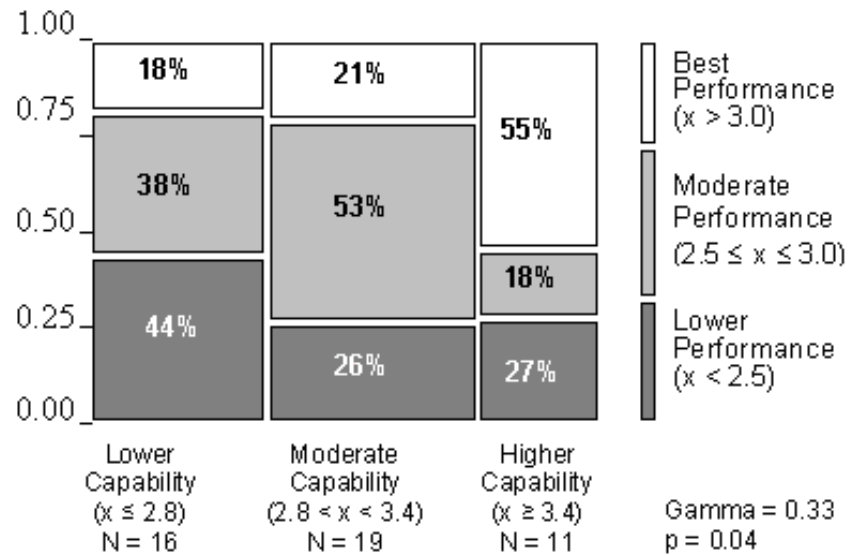


Fig. 10. Relationship Between Requirements Development and Management and Project Performance

For the projects with the lowest Requirements capability, only 18 percent exhibited the best project performance. However, for projects with the highest Requirements capability, 55 percent exhibited the best project performance.

Other SE Categories vs. Project Performance

The impact of each of the 12 aforementioned SE categories upon Project Performance was analyzed in the same manner described in the preceding sections. Additionally, analysis was performed on several ensembles of these SE categories, most notably

- Technical Solution (consisting of the Product Architecture and Trade Study categories) plus Requirements Development and Management
- Overall SE Capability, consisting of all 12 SE categories

Detailed results of these analyses are presented in [1]. Here, we summarize the results solely by the values the strength of relationship with Project Performance, as measured by Gamma. This summary is shown in Fig. 11

Relationship of SE Processes to Program Performance

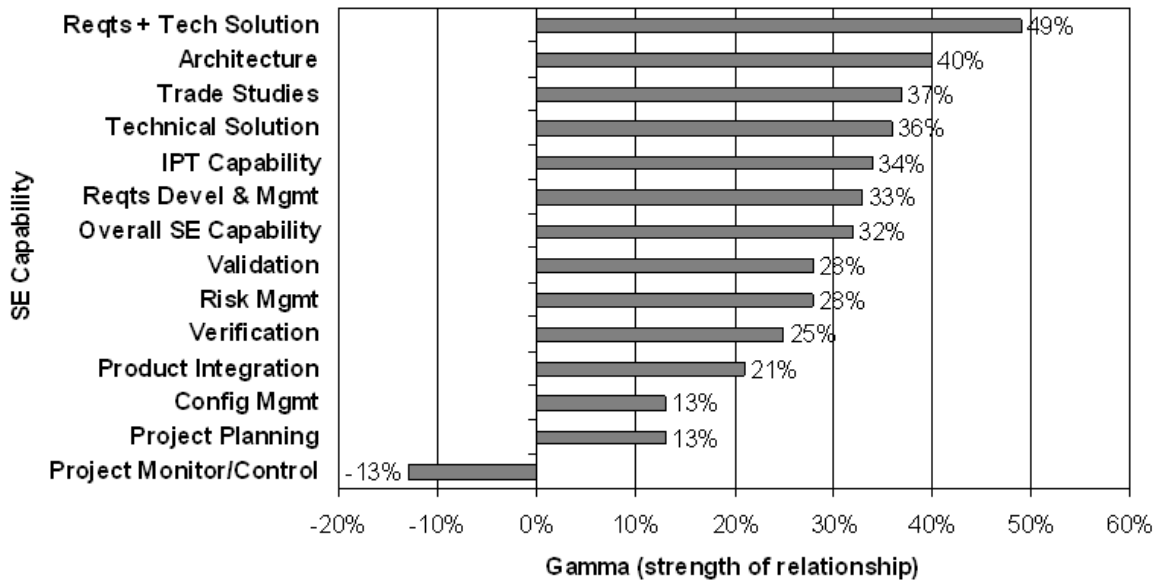


Fig. 11. Summary of Gamma (Relationship Strength) Values for all SE Capability Categories Versus Project Performance

From Fig. 11, note that some of the strongest relationships to Project Performance are seen in the early SE activities (Requirements, Architecture, Trade Studies); the activities typically found on the left side of the Systems Engineering V-model. Also note the negative relationship between Project Monitoring and Control. This does not imply that more Monitoring and Control leads to poorer project performance. Remember that Gamma measures only the strength of the relationship, and does not imply causality. Thus, a more appropriate interpretation of the negative Gamma for Project Monitoring and Control may be “Projects that are performing poorly are subjected to more monitoring and control.”

Since we also measured the level of challenge for each project, we were also able to examine the relationship between:

1. the combined SE capabilities of Technical Solution (comprised of Product Architecture and Trade Studies) and Requirements Development and Management, and
2. Project Performance

using Project Challenge as a control variable. This relationship is shown in Fig. 12. The benefit of these SE capabilities is evident. For projects presenting lower levels of challenge, only 25 percent of those with lower SE capabilities exhibit best performance. However, among these same low-challenge projects, 72 percent of those with higher SE capabilities exhibit best performance. Similarly for projects presenting higher levels of challenge, none (0 percent) of the projects with lower SE capabilities exhibit best performance. However among these same high-challenge projects, 36 percent of those with higher SE capabilities exhibit best performance.

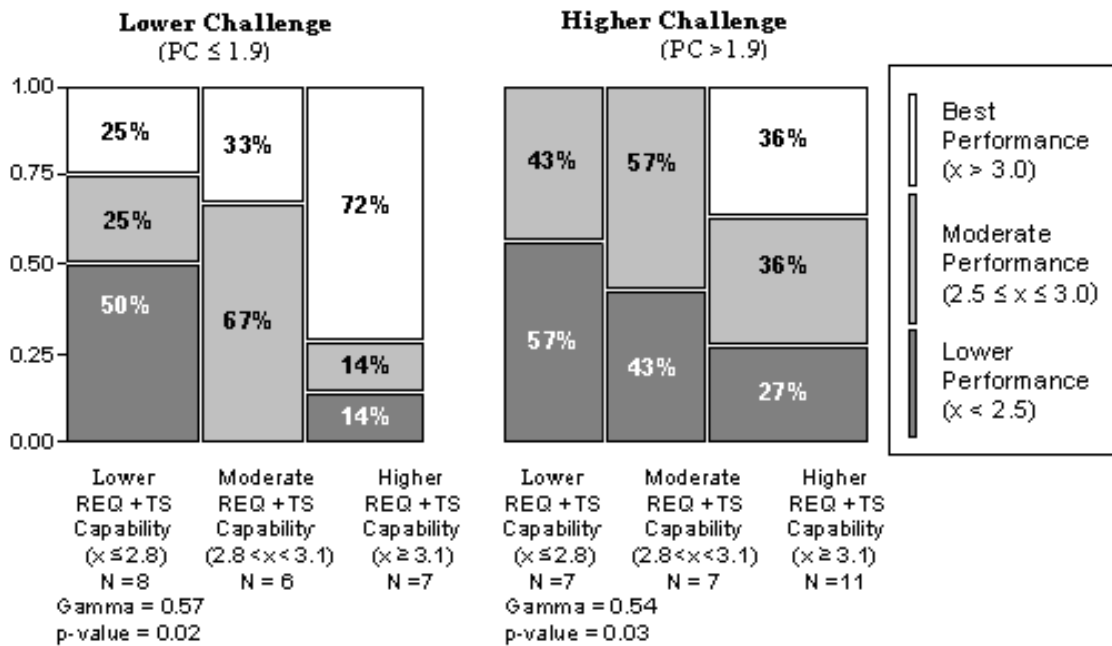


Fig. 12. Control Chart – Requirements + Technical Solution Versus Project Performance Versus Project Challenge

CURRENT AND FUTURE DEVELOPMENTS

The results of the NDIA SE Effectiveness Study have been adopted by several major aerospace and defense suppliers. They have used the survey instrument to assess their internal projects, and have compared the results against the benchmarks established by the NDIA study, using the results to guide SE process improvement activities.

In 2010, the results of this study were briefed to the Director of Systems Engineering at the Office of the Director, Defense Research and Engineering within the Office of the Secretary of Defense (OSD-DDR&E/SE). The findings were positively received, and OSD-DDR&E/SE expressed interest in using findings to improve the value of SE processes within the DoD. He was supportive of a follow-on study to expand the sample size of the study.

Presently, NDIA is collaborating with the IEEE-AESS and SEI to expand and extend the prior study to develop a Business Case for Systems Engineering (BCSE). The mission of this new study is to assist the DoD and the SE community in achieving a quantifiable and persistent improvement in project performance through appropriate application of systems engineering principles and practices. We endeavour to do this by

- Identifying principles and practices PROVEN to provide benefit (by extending and confirming the prior NDIA study)
- Assisting the DoD in developing the guidance to implement the principles and practices identified in study findings
- Assisting the DoD in establishing a means of monitoring / tracking the results of efforts to improve SE
- Assisting DoD in institutionalizing these efforts so that they become “the way DoD does business”.

An overview of the process being used is shown in Figure 13.

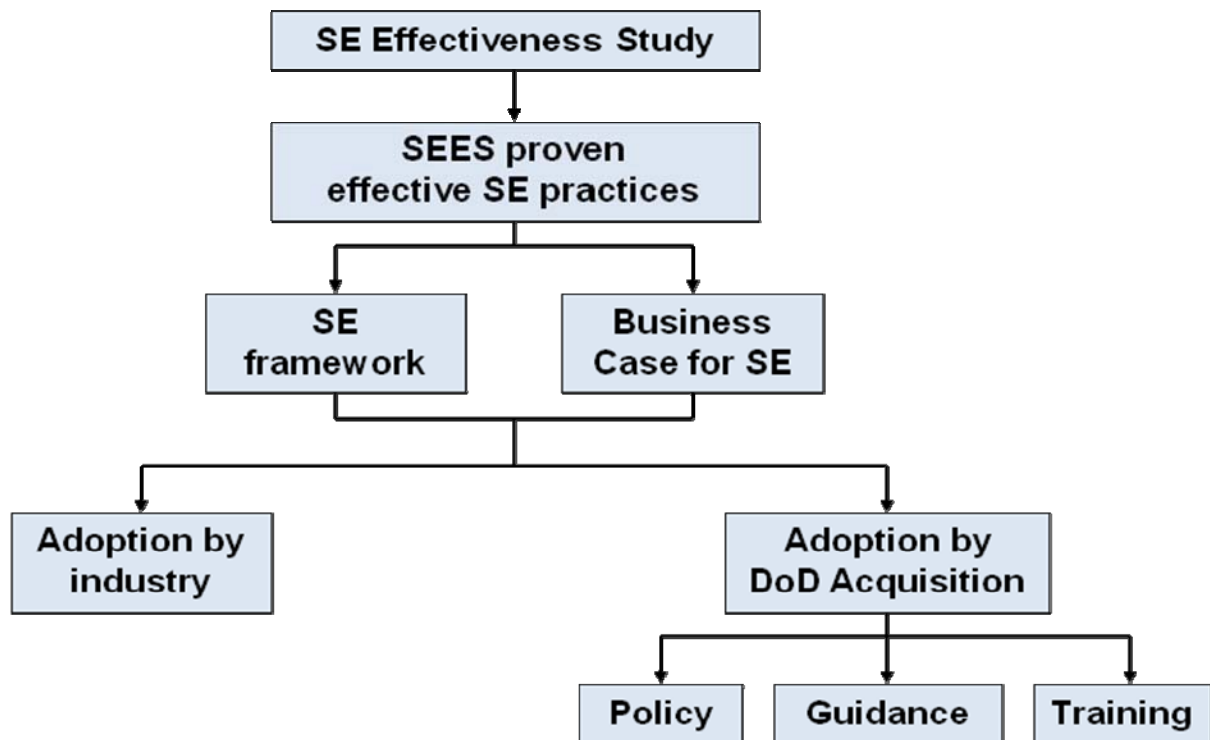


Fig. 13. NDIA / IEEE-AESS / SEI SE EControl Chart – Requirements + Technical Solution vs. Project Performance vs. Project Challenge

This process will be executed in three phases:

- Phase I: Identify SE activities that show net benefit (i.e., value exceeding cost) to program performance. This phase will include the expansion and extension of the previous SE Effectiveness study to survey a larger population, and the development of an SE Framework and a stronger Business Case for SE (BCSE)
- Phase II: Develop recommendations to OSD for policy guidance, and training to implement the findings of Phase I. This phase will include the development of recommendations for OSD review, and the development of tools to implement policy and guidance (e.g., suggested RFP language, suggested contract language, Program Office training, SE assessment methods and tools, SE artifact Data Item Descriptions)
- Phase III: Establish a mechanism for continuous improvement of the BCSE Framework. This phase will include the development of an on-going DoD data collection process integrated with existing program review processes.

Phase I will be completed early in 2012. Phase II will be completed in 2Q2013. Phase III will be completed in 4Q2013.

CONCLUSIONS

The challenge of obtaining consistent measurement of SE application to projects despite the absence of a widely accepted definition of SE has hindered prior efforts of measuring the value of SE. The NDIA SE Effectiveness survey used an innovative method of overcoming this challenge, by focusing on measurement of the application of individual elements of SE (i.e., SE practices) rather than the larger undefined category of SE. This method proved successful, and may have applications in future studies (e.g., the value of software engineering, project

management, process improvement).

This study has provided quantitative evidence of the value of systems engineering activities. The data clearly shows that projects with better SE capabilities have an increased likelihood of delivering better project performance, as measured by budget, schedule and requirements satisfaction. Furthermore, specific categories of SE activities related to Product Architecture, Trade Studies, Requirements Development and Management, and IPT utilization show even stronger relationships with overall project success. The categories of SE activities with the strongest relationships to project success share a few common themes

- they begin early in the project
- they influence the approach to and/or the organization of the project

This knowledge provides a strong incentive for the early application of SE on a project.

Current efforts are proceeding on expanding and extending the earlier study, and implementing the findings within the DoD and the defense community

ACKNOWLEDGMENTS

The author wishes to thank all of the members of the NDIA SE Effectiveness Committee for their years of diligent work that supported this paper. The author particularly expresses appreciation to Al Brown, Bob Bruff, Brian Donahue, Geoff Draper, Khaled El Emam, Dennis Goldenson, Al Mink, Ken Ptack, Bob Rassa, and Mike Uchino. The author also expresses appreciation for the on-going efforts of the joint NDIA – IEEE-AESS SE Effectiveness Committee led by Bill Lyons, Bob Rassa, and Al Brown.

References

1. Elm, J.; Goldenson, D., El Emam, K., Neisa, A., and Donatelli, N. 2008. "A Survey of Systems Engineering Effectiveness". Carnegie Mellon University, Pittsburgh, PA. (available at <http://www.sei.cmu.edu/>)
2. Greuhl, W. 1992. "Lessons Learned, Cost/Schedule Assessment Guide". NASA Comptroller's Office
3. Honour, E. 2004. "Understanding the Value of Systems Engineering".
4. CMMI Product Team. 2001. "Capability Maturity Model[®] Integration (CMMISM), Version 1.1". Carnegie Mellon University, Pittsburgh, PA.

THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

Biography

Joseph Elm is the Project Manager for Strategic Projects within the CERT® program at the Software Engineering Institute at Carnegie Mellon University. He has been engaged in the development of measures of SE effectiveness since 2003. Mr. Elm's career has included engineering, project management, technical management, and executive management positions in companies within aerospace, defense, and transportation industries. Mr. Elm holds a BSEE degree from Carnegie Mellon University and an MSEE degree from the University of Pittsburgh.