Software Engineering Institute | Carnegie Mellon University

Quantifying the Effectiveness of Systems Engineering

Presenters: Joseph P. Elm

The Software Engineering Institute (SEI)
a DoD Research FFRDC

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM-0001639

The Software Engineering Institute (SEI)

Who we are

- A Federally Funded Research and Development Center (FFRDC)
- Sponsored by the Department of Defense
  - but we work for all government agencies
- Created in 1984 / CERT program founded in 1988
- A Part of Carnegie Mellon University

What we do

- Software Engineering
- Software Research
- Cybersecurity
- Assurance
- Acquisition Solutions
- Emerging Technologies

Our Mission

To advance the technologies and practices needed to acquire, develop, operate, and sustain software systems that are innovative, affordable, trustworthy, and enduring.

SEI R&D Technical Priorities and Goals

Major Areas	SEI Technical Priorities	Goals Include . . .
Software Engineering	Development methods, empirical analysis methods, cost estimation, validation, sustainment	Create and sustain affordable, trustworthy, effective and enduring software systems with acceptable urgency
Assurance	Designed-in security, evidence, acquisition guidance, tools	Improve the level of assurance in software systems using evidence
Specific Capabilities	Algorithms, networks and networking, mobile applications, embedded/real-time distributed systems	Maintain and expand the toolbox of techniques in critical, emerging and pervasive areas
Cybersecurity	Forensics, coding standards, insider threat behavior, malware and code analysis, workforce education	Improve base and operational security, understand adversaries, spreading cyber competence

What does it take to develop a complex system?

Many Systems

- Propulsion
- Hydraulics
- Power
- Controls
- Radar
- Structures
- Navigation
- Computers
- Communications
- …

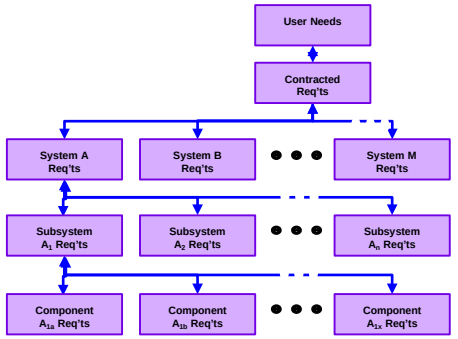
Many disciplines

- Mechanical Engineering
  - fluidynamics
  - structural
- Metallurgical Engineering
- Electrical Engineering
  - power
  - radar
  - Communications
- Manufacturing Engineering
- Software Engineering
- Test Engineering
- …

But, Not Everything Fits Cleanly into One Discipline

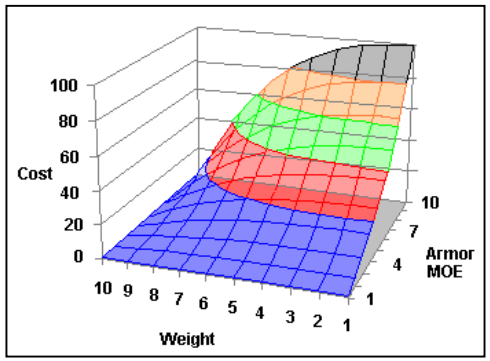
Requirements Development and Management

- Decomposition of requirements
- Allocation of requirements among multiple systems



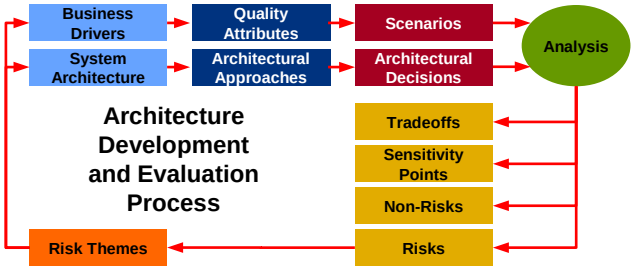
Interdisciplinary Trade Studies

- Requirements implementation in hardware vs. software
- Exotic alloys for low weight vs. more common materials for low cost
- Lower radar cross section vs. higher aerodynamic performance



Architecture Development

- Model Driven Design
- Quality Attribute Driven Architecture



Who Pulls it All Together ?

The Systems Engineer

Required skills

- Global system-wide perspective
- Full life-cycle perspective
- Forward-looking
- Multidisciplinary technical knowledge
- Fact-based decision-making
- Multi-tasking

Tasks Performed *

- Requirements Development
- Requirements Management
- Trade Studies
- System Architecture Development
- Interface Management
- Configuration Management
- Program Planning
- Program Monitoring and Control
- Risk Management
- Product Integration Planning and Oversight
- Verification Planning and Oversight
- Validation Planning and Oversight

How likely is program success if these activities are not done well?

* Some tasks are done in partnership with the Program Manager

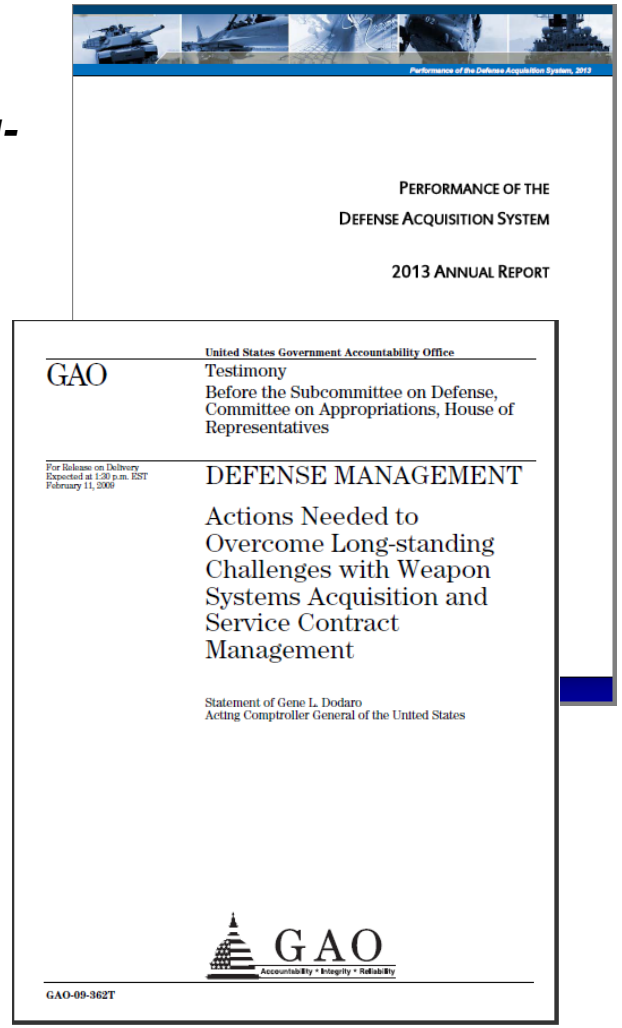
Challenges in DoD Acquisition

GAO-09-362T - Actions Needed to Overcome Long-standing Challenges with Weapon Systems Acquisition and Service Contract Management

- "costs … increased 26% and development costs increased by 40% from first estimates"
- "programs … failed to deliver capabilities when promised —often forcing warfighters to [maintain] legacy systems"
- "current programs experienced, on average, a 21-month delay in delivering initial capabilities to the warfighter"

Although DoD is the largest acquirer in the world, acquisition troubles remain [1]

- 2011 MDAP RDT&E cost growth (mean) 84%
- 2011 MDAP Procurement cost growth (mean) 28%
- Effectiveness (1984-2011) 89%
- Suitability (1984-2011) 72%
- Nunn-McCurdy breach rate from 1997-2011 31%



1. "Performance of the Defense Acquisition System 2013 Annual Report" Table 2-3, page 34)

Root Cause of Poor Program Performance

Inadequate Systems Engineering!

- Finding from *Performance of the Defense Acquisition System 2013 Annual Report*
  - **Dominant root cause** of MDAP Cost Growth
- Finding from *GAO-09-362T*
  - "… managers rely heavily on assumptions about system requirements, technology, and design maturity, which are consistently too optimistic. These gaps are largely the result of a **lack of a disciplined systems engineering analysis** prior to beginning system development …"

MDAP Cost Growth: PARCA Root Cause Analysis[1]

	Dominant
10 of 18 (56%)	**Poor management performance** • **Systems engineering** • Contractual incentives • **Risk management** • Situational Awareness
5 of 18 (28%)	**Baseline cost and schedule estimates** • **Framing assumptions**
4 of 18 (22%)	**Change in procurement quantity**
	Infrequent
1 of 18	**Immature technology, excessive manufacturing, or integration risk**
2 of 18	**Unrealistic performance expectations**
1 of 18	**Unanticipated design, engineering, manufacturing or technology issues**
None	**Funding inadequacy**

1. "Performance of the Defense Acquisition System 2013 Annual Report" Table 2-3, page 34)

Perceptions of SE

The SE efforts on my program are critical because they …

- … pay off in the end.
- … ensure that stakeholder requirements are identified and addressed.
- … provide a way to manage program risks.
- … establish the foundation for all other aspects of the design.
- … optimize the design through evaluation of alternate solutions.

We need to minimize the SE efforts on this program because …

- … including SE costs in our bid will make it non-competitive.
- … we don't have time for '*paralysis by analysis*.' We need to get the design started.
- … we don't have the budget or the people to support these efforts.
- … SE doesn't produce deliverable outputs.
- … our customer won't pay for them.

These are the ASSERTIONS, but what are the FACTS?

What is the ROI for SE?

It's difficult to justify the costs of SE in terms that program managers and corporate managers can relate to.

- The costs of SE are evident
  - Cost of resources
  - Schedule time
- The benefits are less obvious and less tangible
  - Cost avoidance (e.g., reduction of rework from interface mismatches)
  - Risk avoidance (e.g., early risk identification and mitigation)
  - Improved efficiency (e.g., clearer organizational boundaries and interfaces)
  - Better products (e.g., better understanding and satisfaction of stakeholder needs)`

We need to quantify the effectiveness and value of SE by examining its effect on program performance?

Measuring ROI

Obtain quantitative evidence of the costs and associated benefits of Systems Engineering activities via a survey of development programs

The SE Effectiveness Study

Purpose

- Strengthen the business case for SE by relating program performance to the use of SE practices.

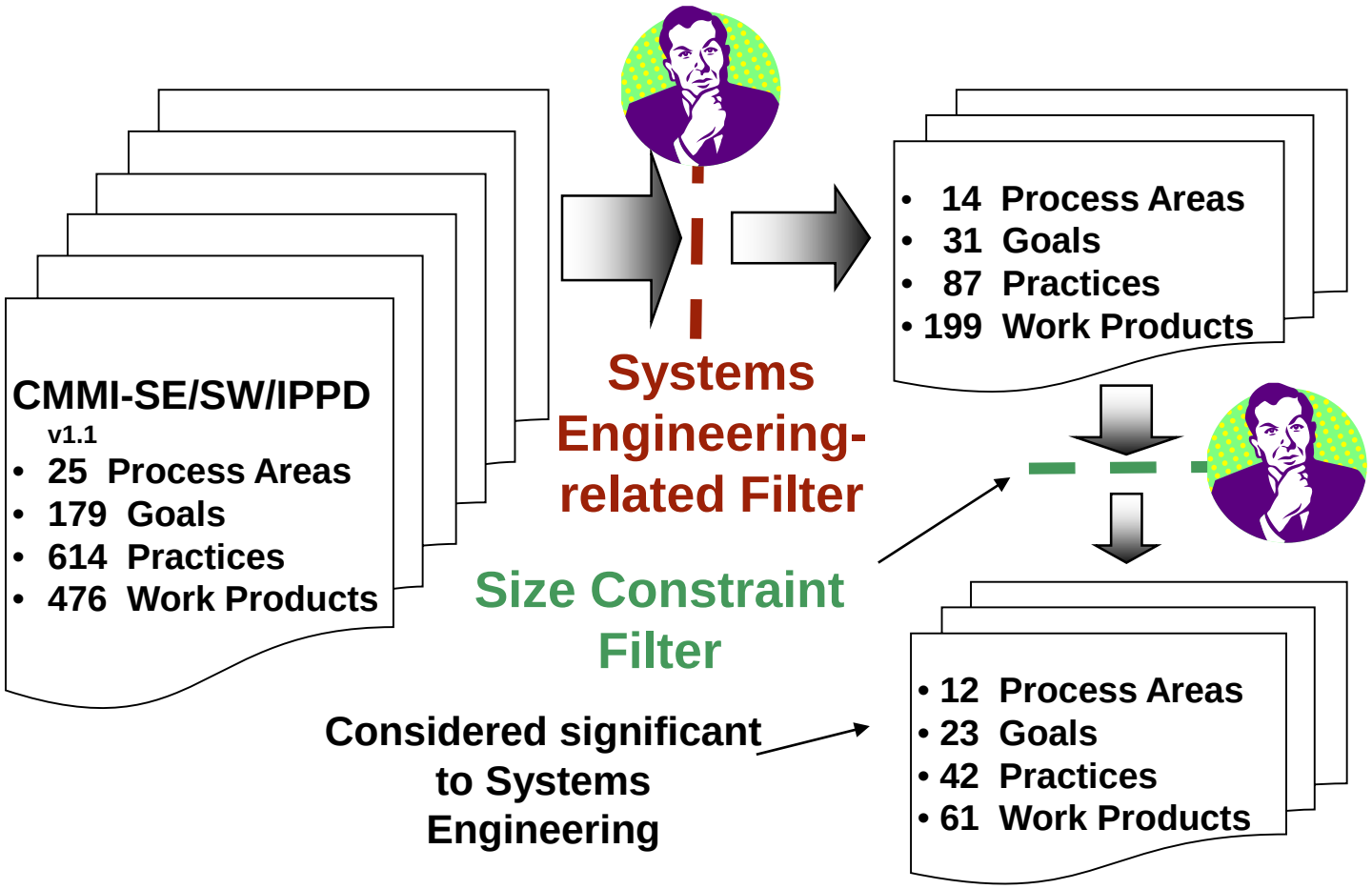
Method

- Contact development programs using the resources of NDIA, AESS, and INCOSE.
- Survey programs to assess their
  - SE activities
  - Program performance
  - Degree of challenge
- Process responses to identify statistical relationships between parameters.

Survey Tenets

- All data is submitted anonymously and handled confidentially by the SEI.
- Only aggregated non-attributable data is released.

Artifact-based assessment of SE Practices



Survey content is based on a recognized standard (CMMI)

Assessment of Program Performance

Assess TOTAL Program Performance

- Program Cost, Program Schedule, Technical Performance
- Focus on commonly used measurements
  - EVMS, baseline management
  - requirements satisfaction
  - budget re-baselining and growth
  - milestone and delivery satisfaction

Assessment of Other Factors

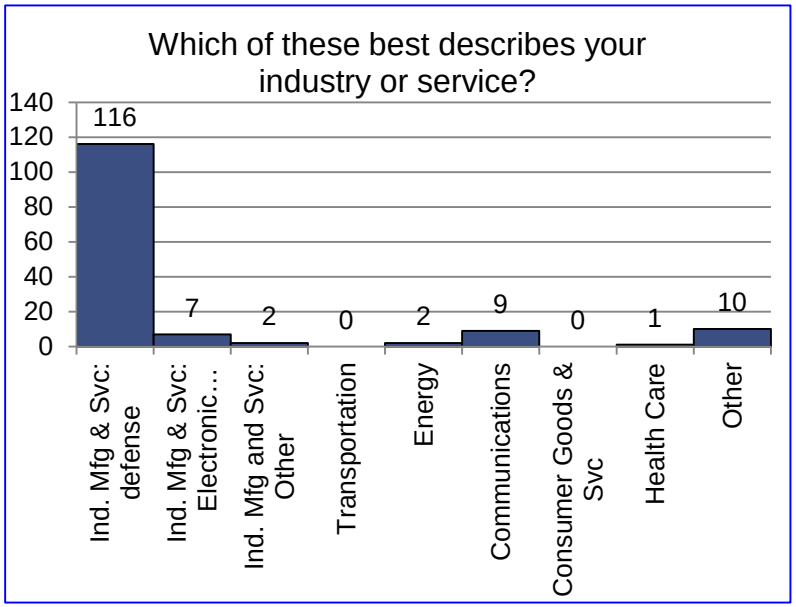
- **Program Challenge** – some programs are more complex than others
- **Prior Experience** – some acquirers are more capable than others

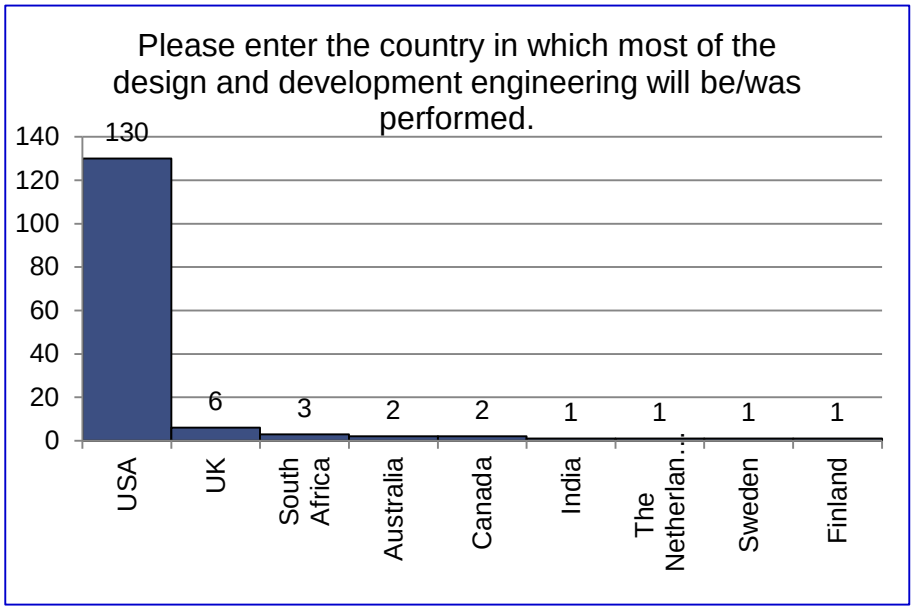
Study Participants

Participant Solicitation

- Contacted key members of major defense contractors to promote study participation
- Contacted the memberships of NDIA SE Division, IEEE AESS, and INCOSE

Collected 148 valid responses

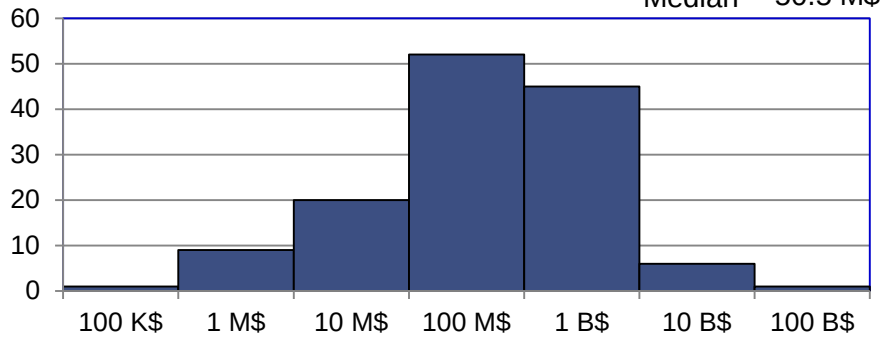




Study Results

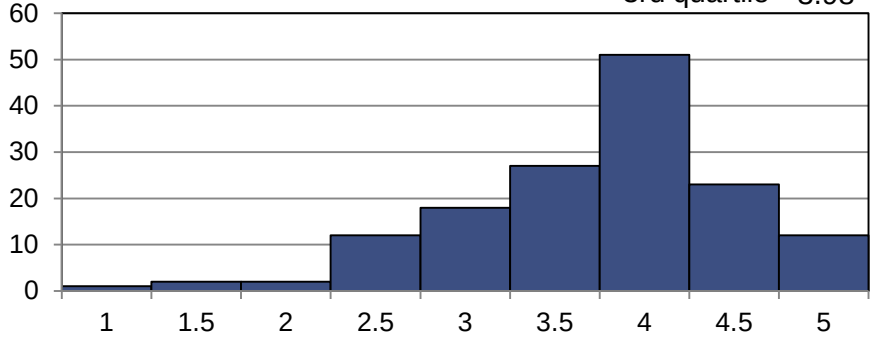
Total program contract value

Mean	488 M$
Std. Dev.	2.22 B$
Median	50.5 M$



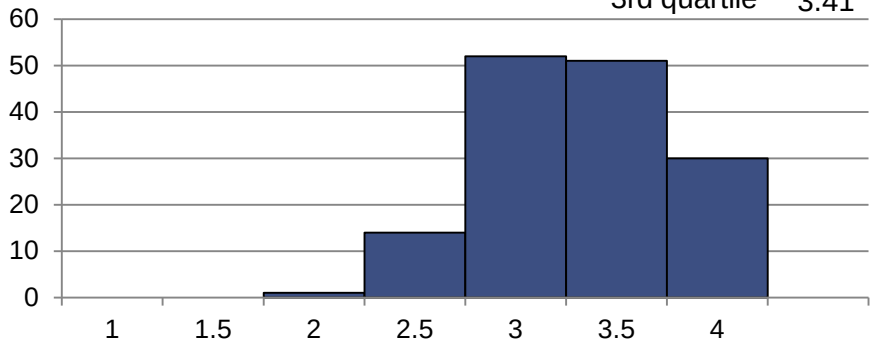
Program Performance (Perf)

1st quartile	3.03
Median (2nd quartile)	3.58
3rd quartile	3.98



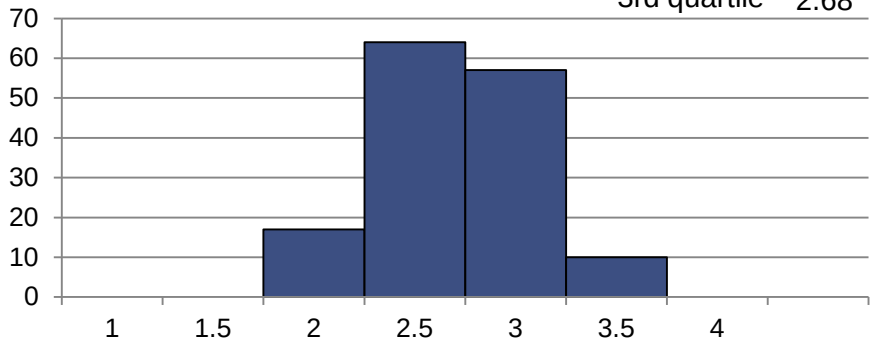
Total SE Deployed on Program (SEC_Total)

1st quartile	2.78
Median (2nd quartile)	3.03
3rd quartile	3.41

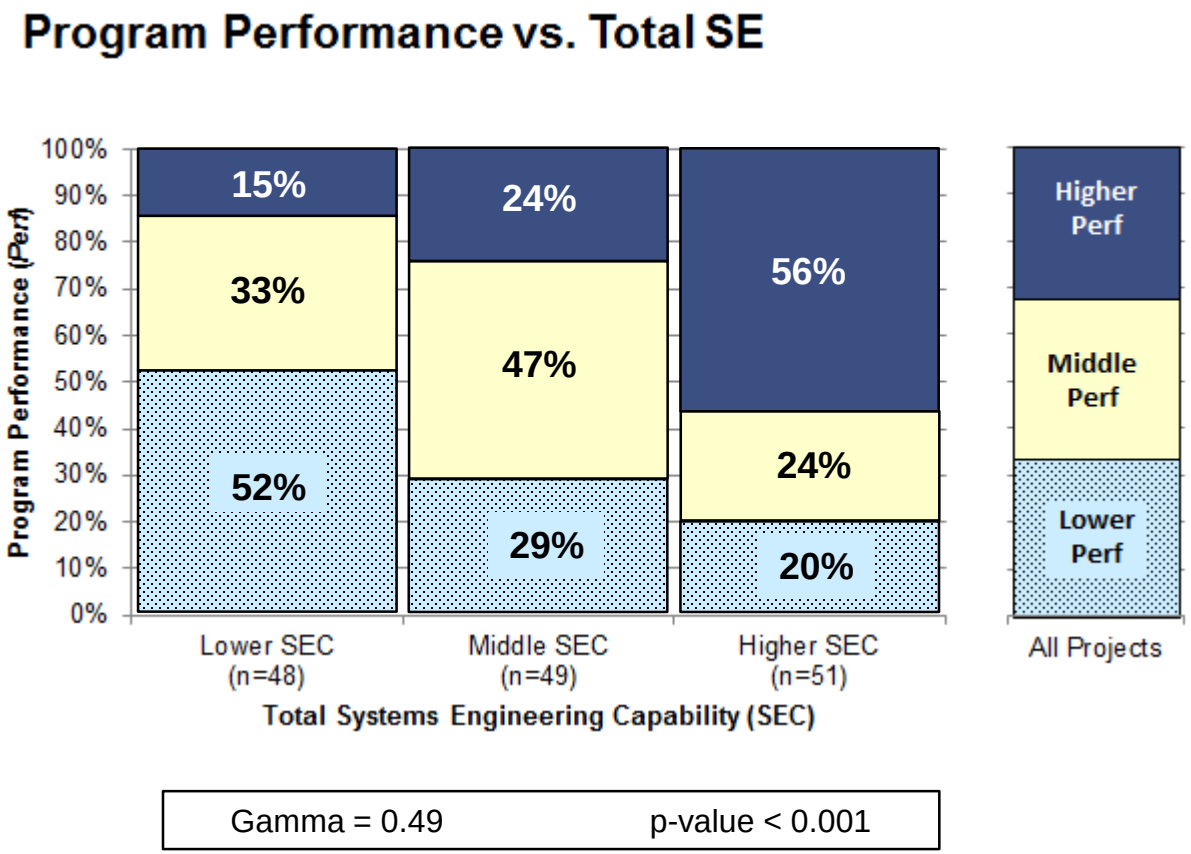


Program Challenge (PC)

1st quartile	2.22
Median (2nd quartile)	2.50
3rd quartile	2.68



The Bottom Line: SE = Performance



Across ALL programs, 1/3 are at each performance level

For Lower SEC programs, only 15% deliver higher performance

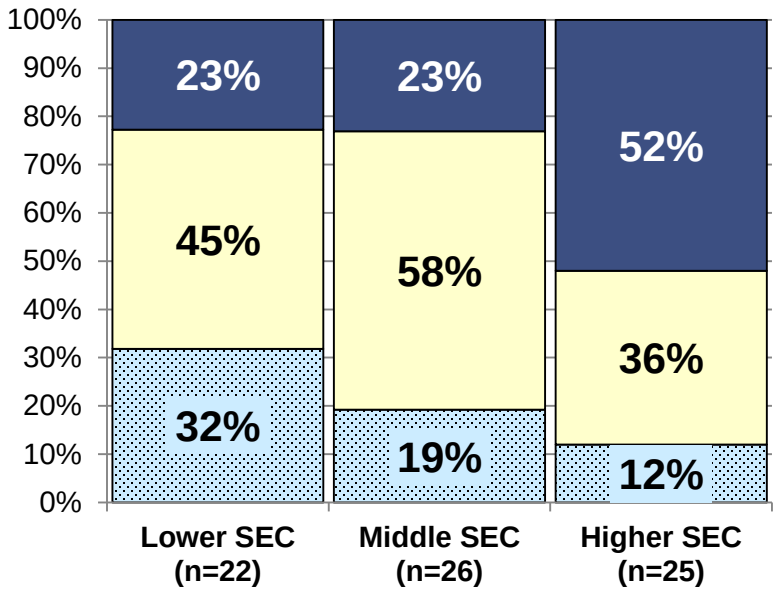
For Middle SEC programs, 24% deliver higher performance

For Higher SEC programs, 57% deliver higher performance

Gamma = 0.49 represents a VERY STRONG relationship

For Challenging Programs SE is even MORE important

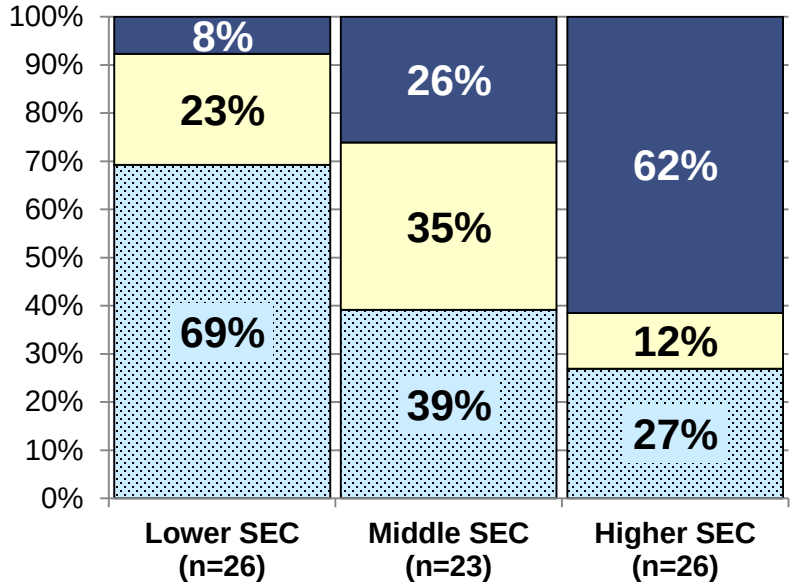
Perf vs. SEC_Total (Low PC)



Gamma = 0.34 p-value = 0.029

A STRONG relationship between Total SE and Program Performance for LOWER CHALLENGE programs

Perf vs. SEC_Total (High PC)



Gamma = 0.62 p-value = 0.000

A VERY STRONG relationship between Total SE and Program Performance for HIGHER CHALLENGE programs

A Deeper Look at SE Activities

Our survey questions addressed 11 areas of SE Activities

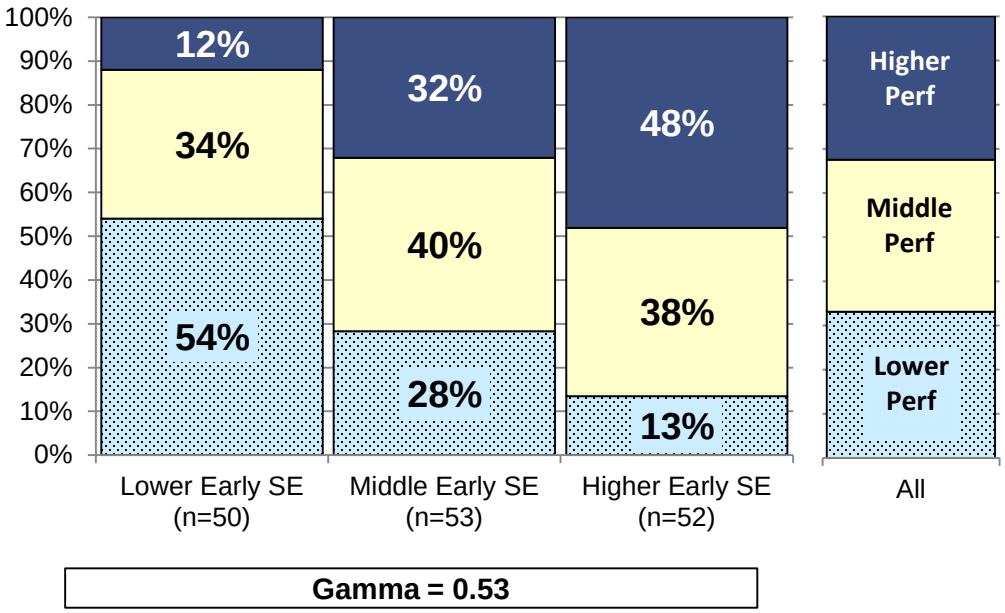
- Program Planning
- Requirements Development and Management
- Product Architecture
- Trade Studies
- Product Integration
- Verification
- Validation
- Risk Management
- Configuration Management
- Integrated Product Teams
- Program Monitoring and Control

Early SE

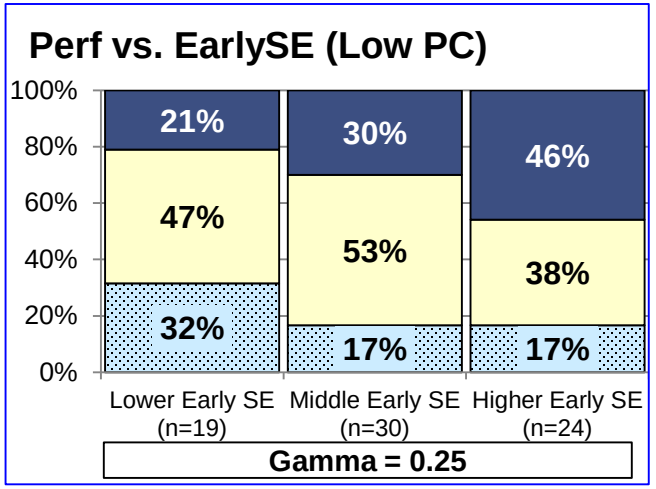
This enabled us to assess a program's deployment of SE in each of these areas

Early SE is the MOST Important

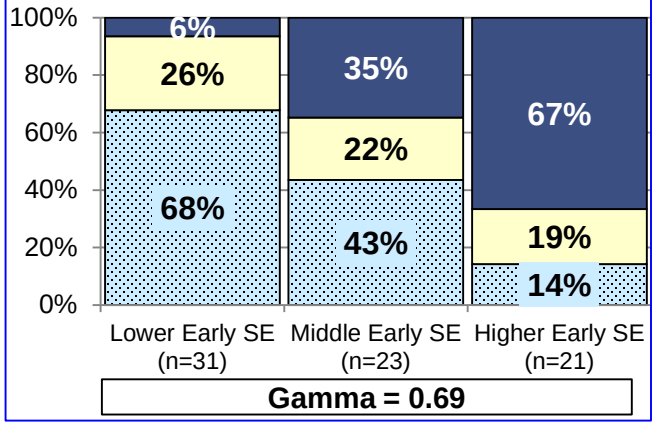
Perf vs. EarlySE



Perf vs. EarlySE (Low PC)



Perf vs. EarlySE (High PC)

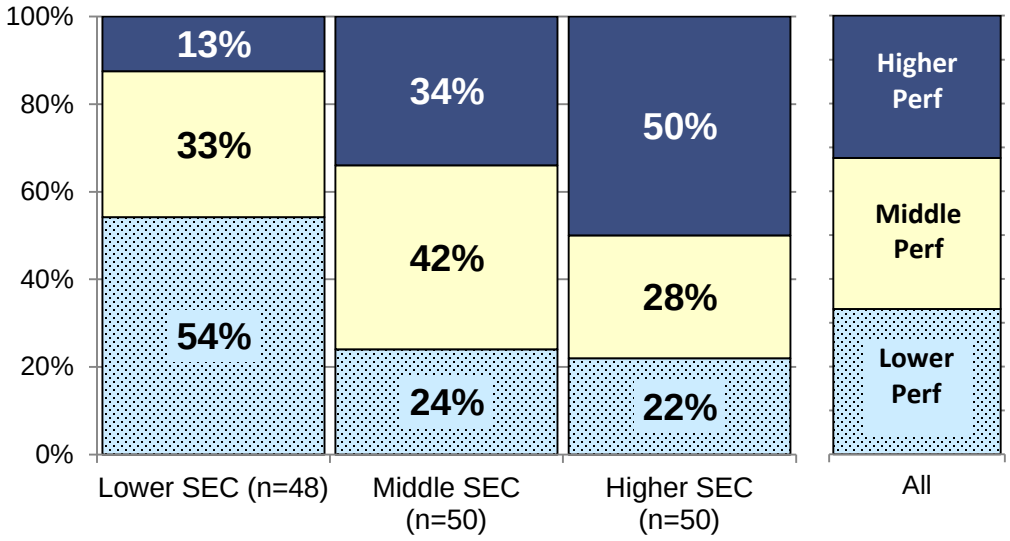


The relationship:

for the set of all programs	**0.53 = Very Strong**
for the set of Low Challenge programs	**0.25 = Moderate**
for the set of High Challenge programs	**0.69 = Very Strong**

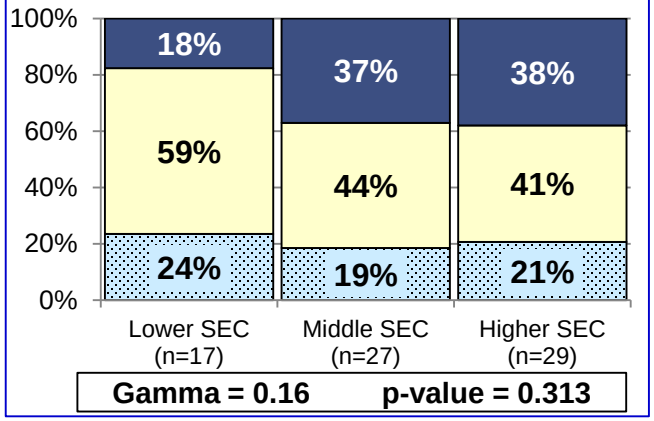
Program Planning vs. Performance

Perf vs. SEC-PP



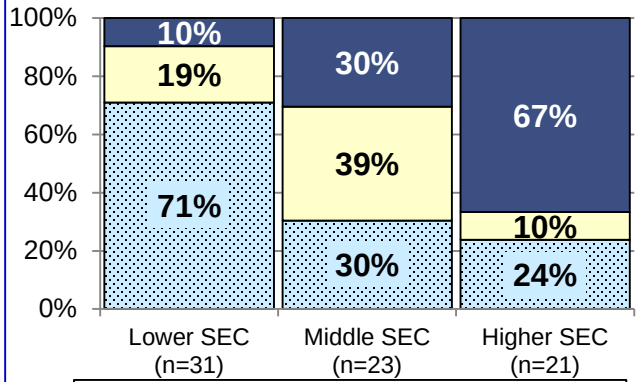
Gamma = 0.46 p-value = 0.000

Perf vs. SEC-PP (Low PC)



Gamma = 0.16 p-value = 0.313

Perf vs. SEC-PP (High PC)



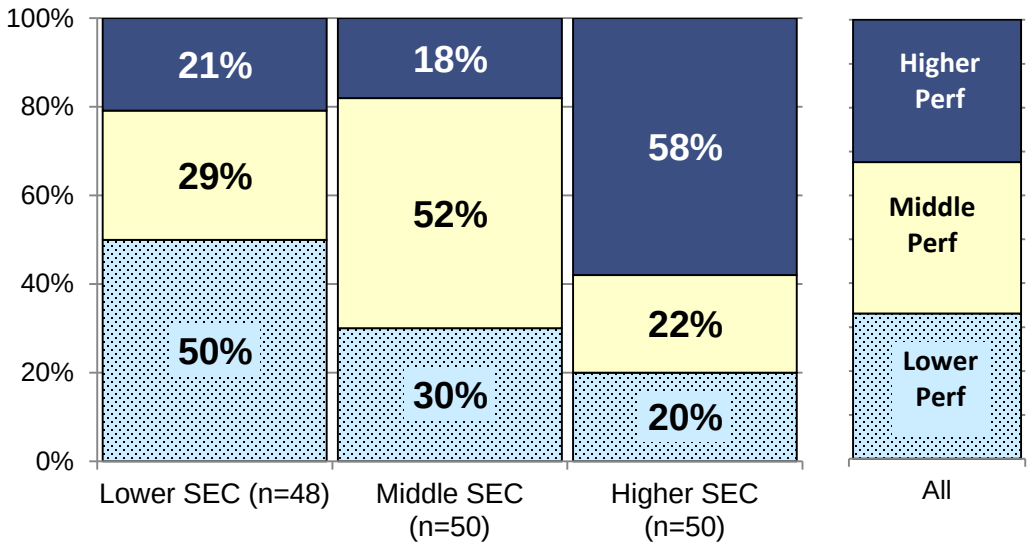
Gamma = 0.65 p-value = 0.000

The relationship:

for the set of all programs	**0.46 = Very Strong**
for the set of Low Challenge programs	**0.16 = Weak**
for the set of High Challenge programs	**0.65 = Very Strong**

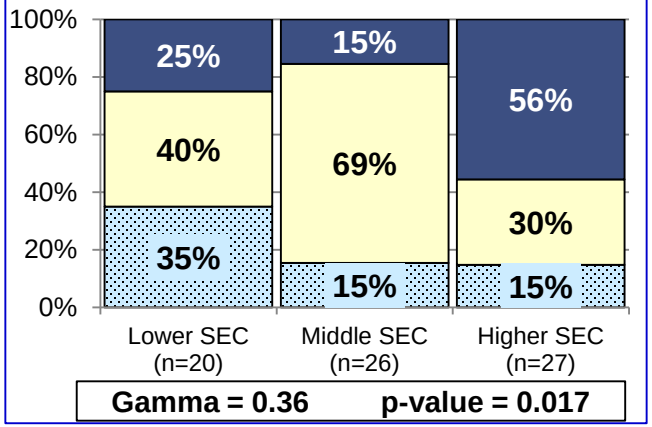
Requirements Dev't & Mg't vs. Performance

Perf vs. SEC-REQ



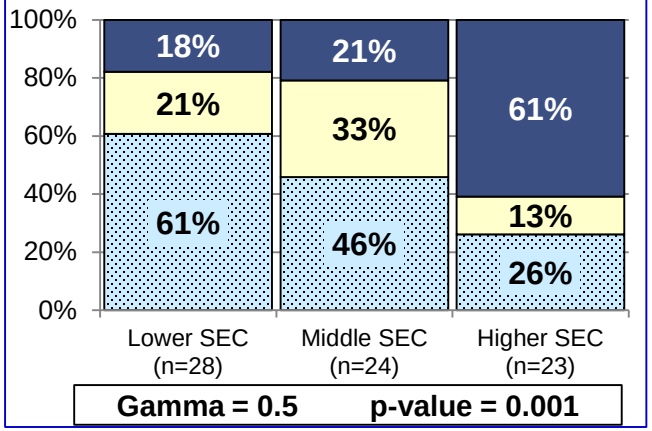
Gamma = 0.44 **p-value = 0.000**

Perf vs. SEC-REQ (Low PC)



Gamma = 0.36 **p-value = 0.017**

Perf vs. SEC-REQ (High PC)



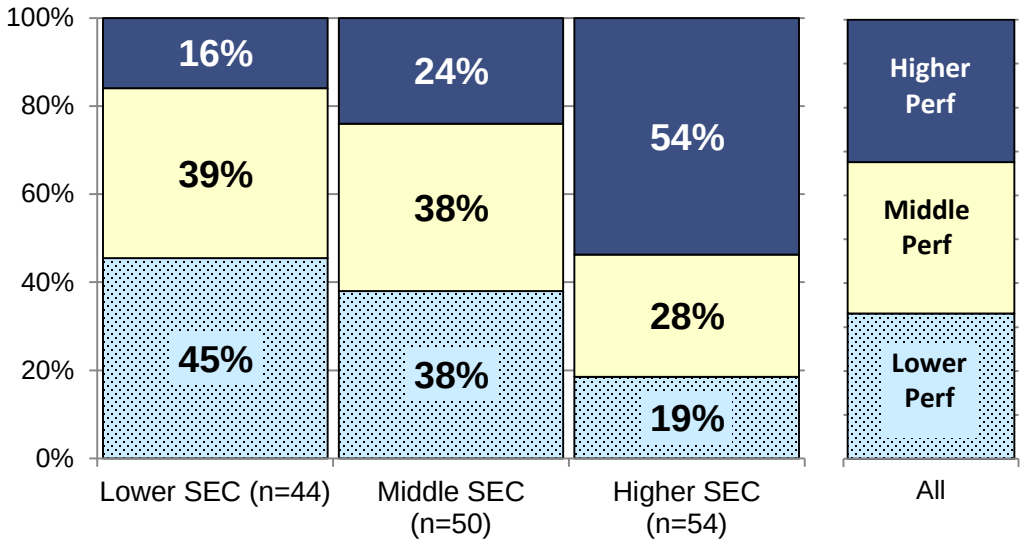
Gamma = 0.5 **p-value = 0.001**

The relationship:

for the set of all programs	**0.44 = Very Strong**
for the set of Low Challenge programs	**0.36 = Strong**
for the set of High Challenge programs	**0.50 = Very Strong**

Verification vs. Performance

Perf vs. SEC-VER

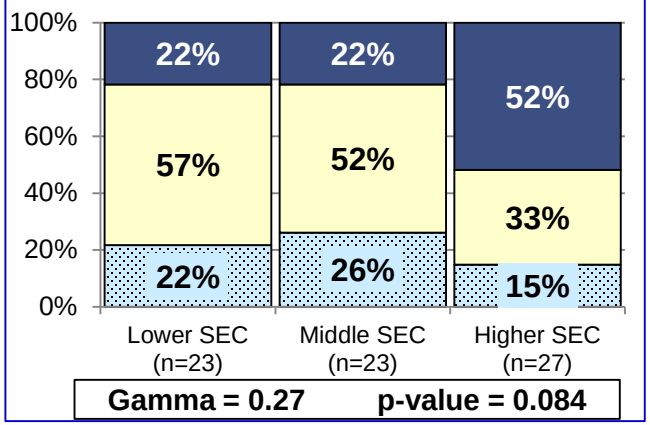


Gamma = 0.43 p-value = 0.000

The relationship:

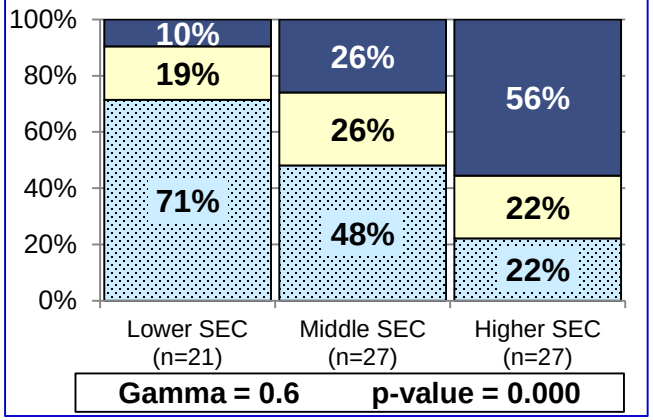
for the set of all programs	**0.43 = Very Strong**
for the set of Low Challenge programs	**0.27 = Moderate**
for the set of High Challenge programs	**0.60 = Very Strong**

Perf vs. SEC-VER (Low PC)



Gamma = 0.27 p-value = 0.084

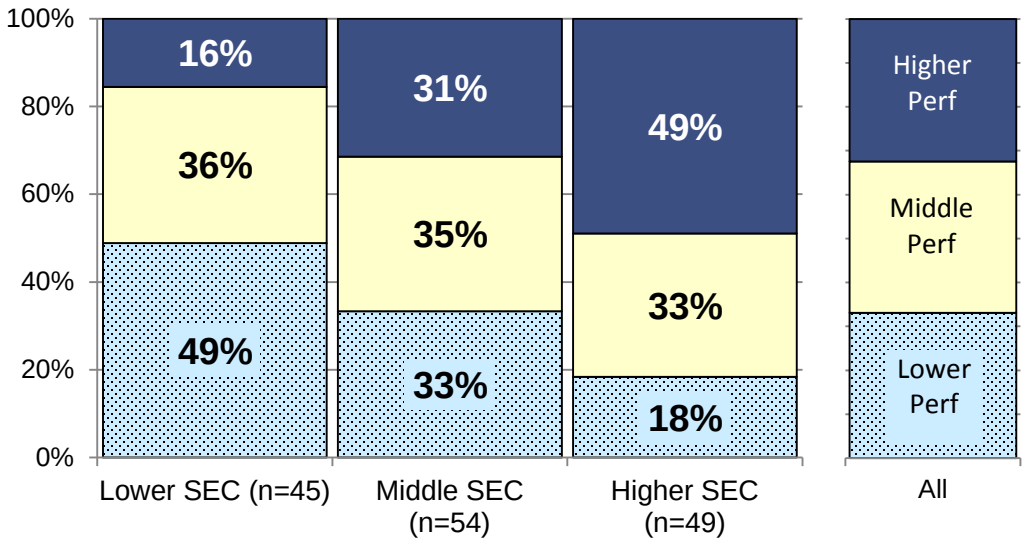
Perf vs. SEC-VER (High PC)



Gamma = 0.6 p-value = 0.000

Architecture vs. Performance

Perf vs. SEC-ARCH

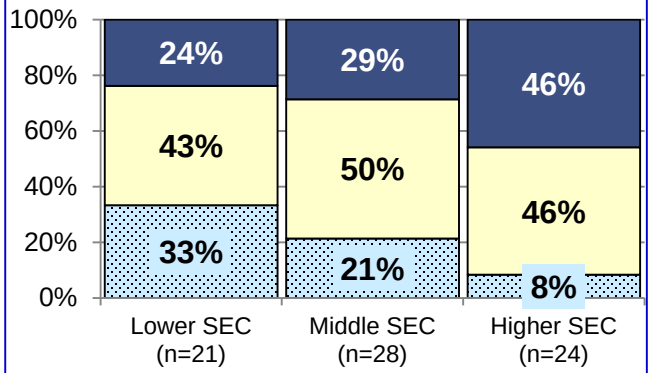


Gamma = 0.41 p-value = 0.000

The relationship:

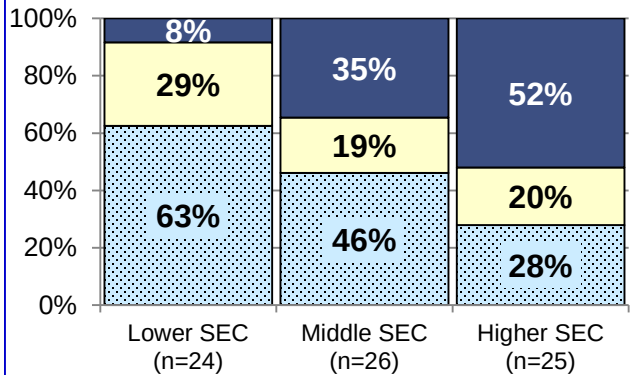
for the set of all programs	**0.41 = Very Strong**
for the set of Low Challenge programs	**0.31 = Strong**
for the set of High Challenge programs	**0.49 = Very Strong**

Perf vs. SEC-ARCH (Low PC)



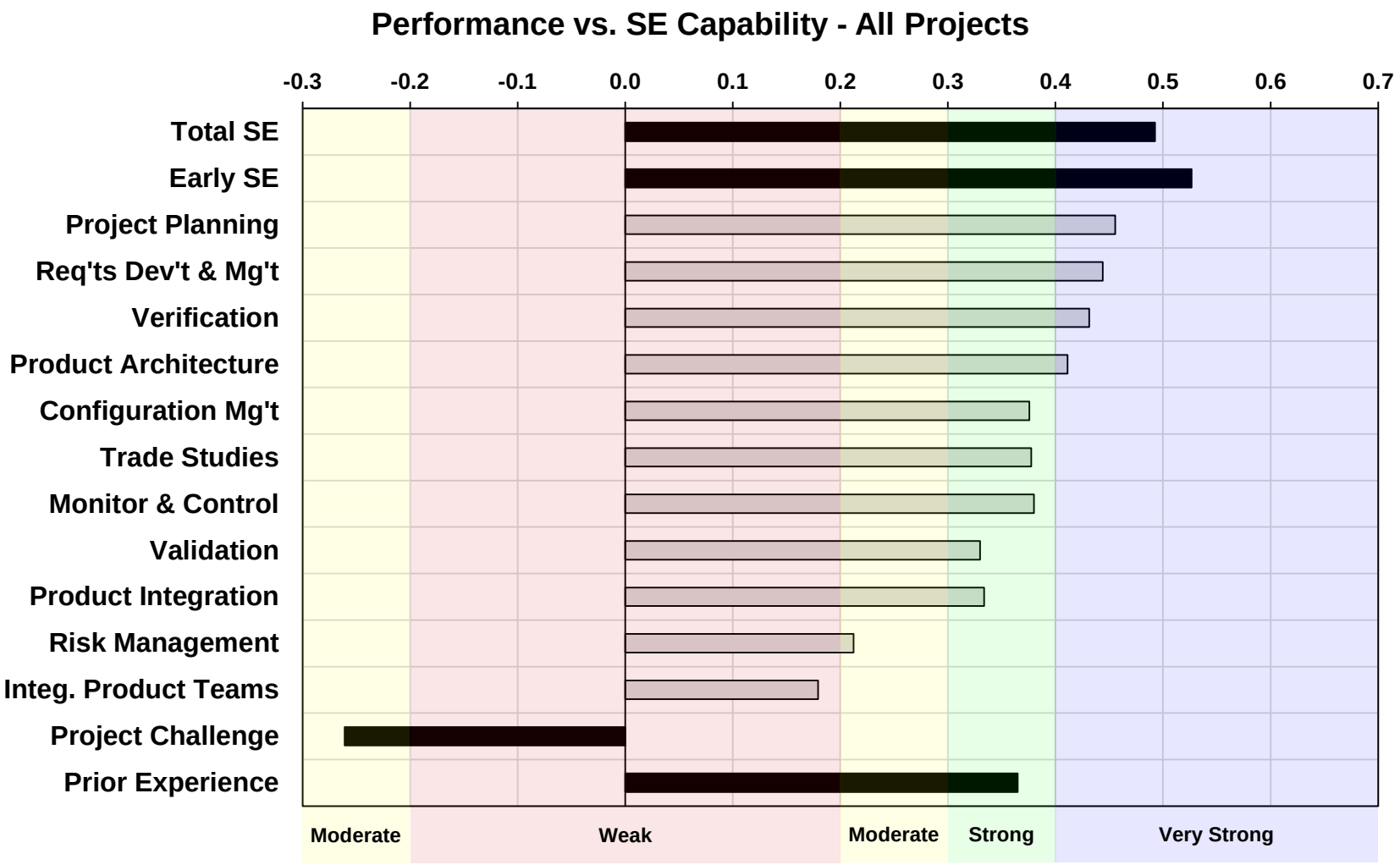
Gamma = 0.31 p-value = 0.051

Perf vs. SEC-ARCH (High PC)

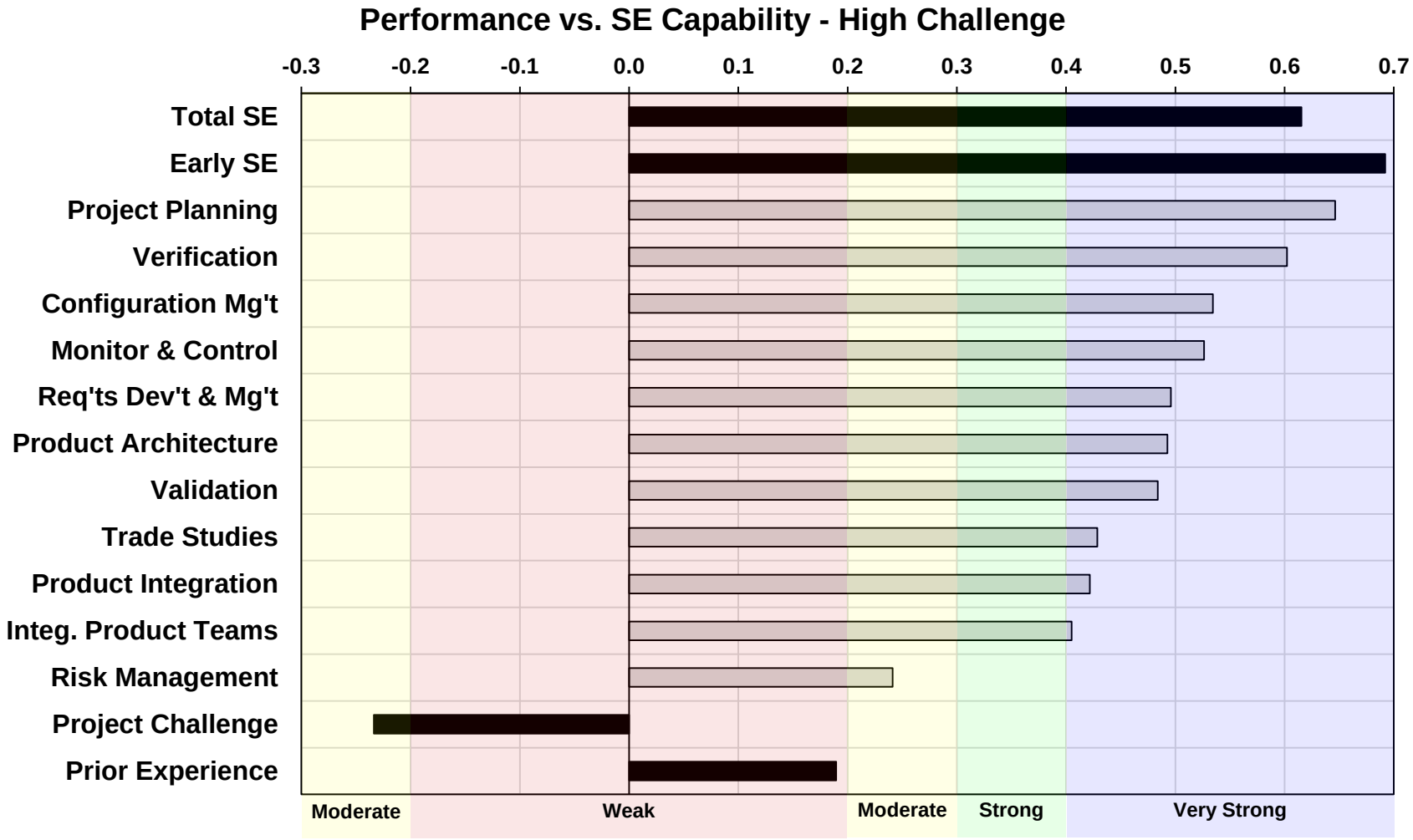


Gamma = 0.49 p-value = 0.001

Summary of Relationships – All Projects



Summary of Relationships - Challenging Projects

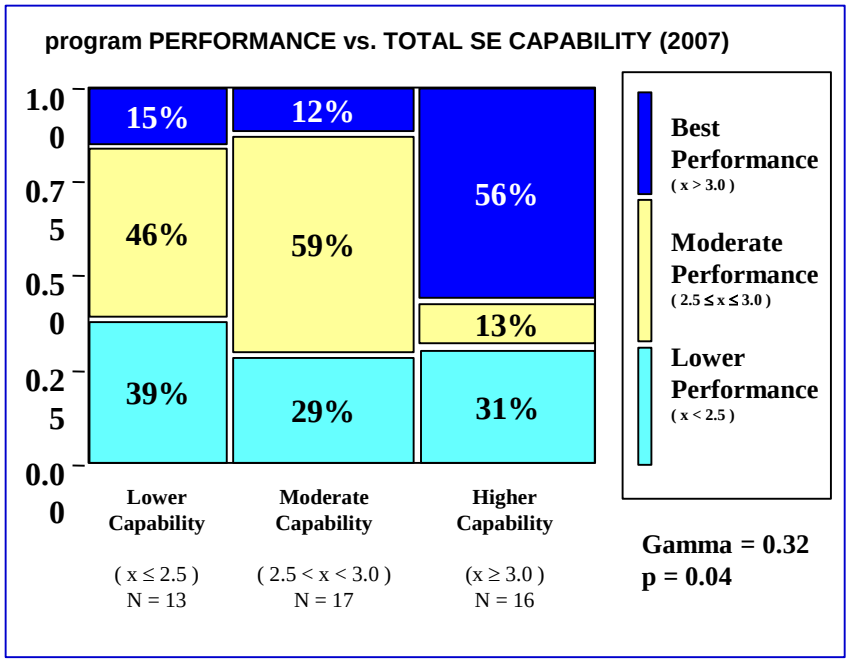


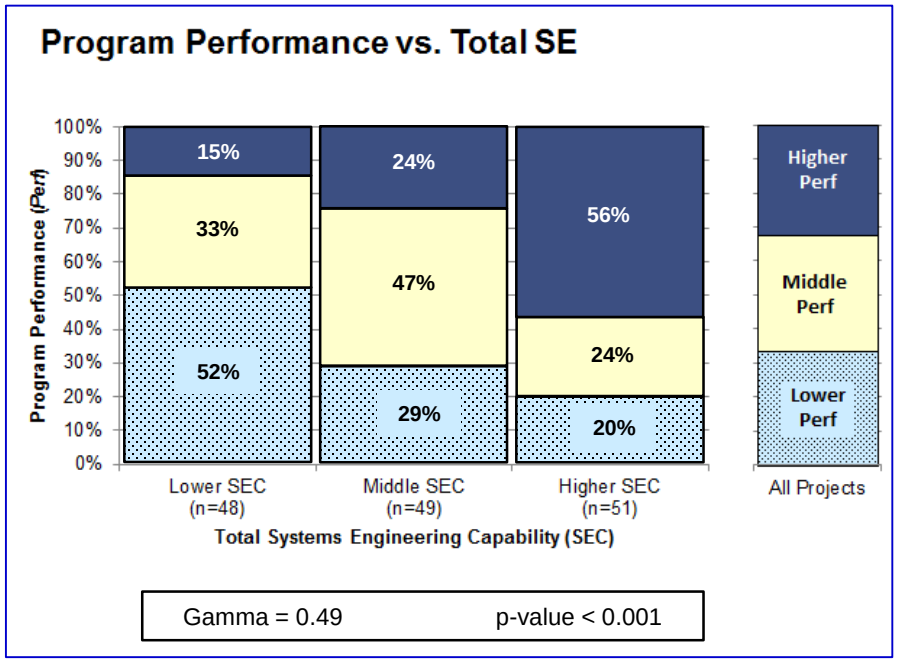
Comparison with 2007 SE Effectiveness Study

On the whole, relationships identified in this study are noticeably stronger than those from the previous study

- Probably due to reduction in noise resulting from the larger sample size

Most results from the two studies are generally in agreement





Using the Findings of This Study -1

System Developers can use this report to:

- plan SE capability improvement efforts focusing on those SE activities most strongly associated with improved program performance
- serve as an industry benchmark for their organization's SE performance.
  - Assess programs within the organization and compare with the study results to leverage strengths, and improve weaknesses
- justify and defend SE activities applied to programs.

System Acquirers may use this report to:

- incorporate SE requirements into RFPs and source selection activities
  - Ensure that SE activities are included in schedules and budgets
  - Demand SE deliverables (e.g. SE Management Plan) during program execution
  - Require SE evaluations of contractors during source selection and during program execution
- employ this survey or similar methods to collect data from during program execution as a means of identifying supplier SE deficiencies contributing to program risks.

Using the Findings of This Study -2

SE Educators may use this report to:

- Focus curricula on key aspects of SE
- Convey to students the value of SE

All may use this report to:

- identify critical SE capabilities to guide Workforce Development

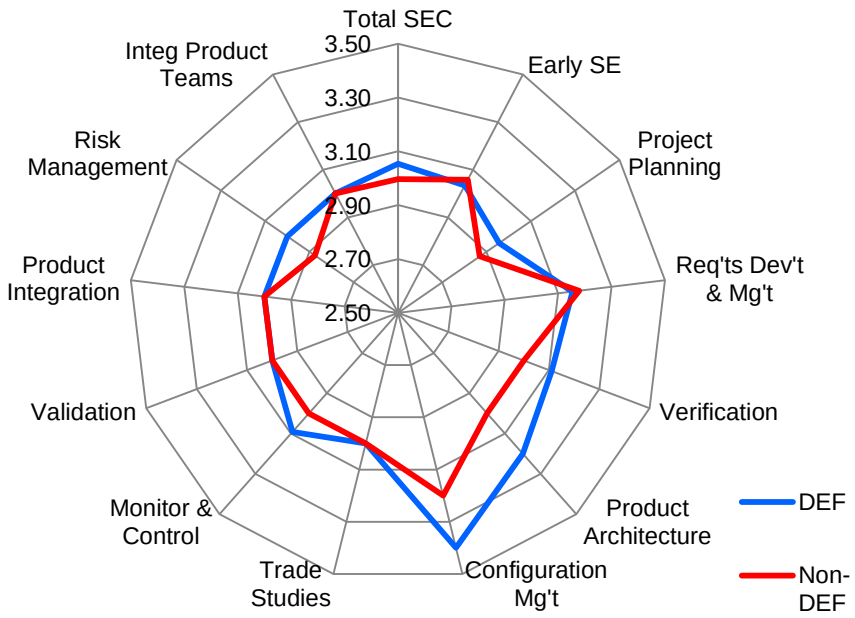
Defense vs. Non-defense Projects -1

NEW !

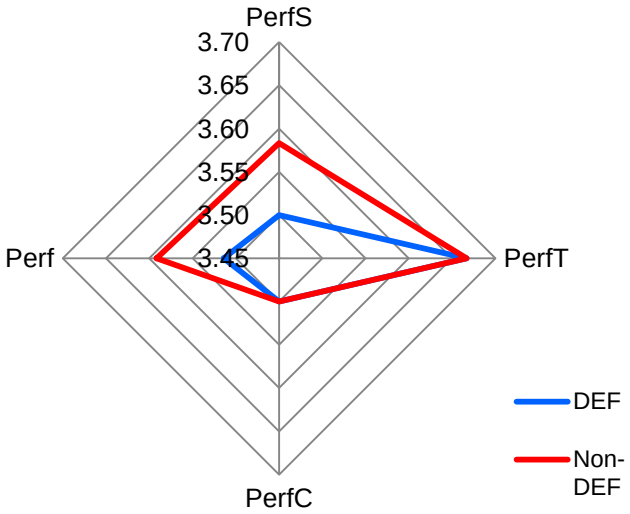
The data from the 2012 SE Effectiveness Study included responses from both defense-domain and non-defense-domain projects

- Cross-domain comparison of SE deployment, project performance and the relationships between them can identify improvement opportunities through transplantation of best practices between domains

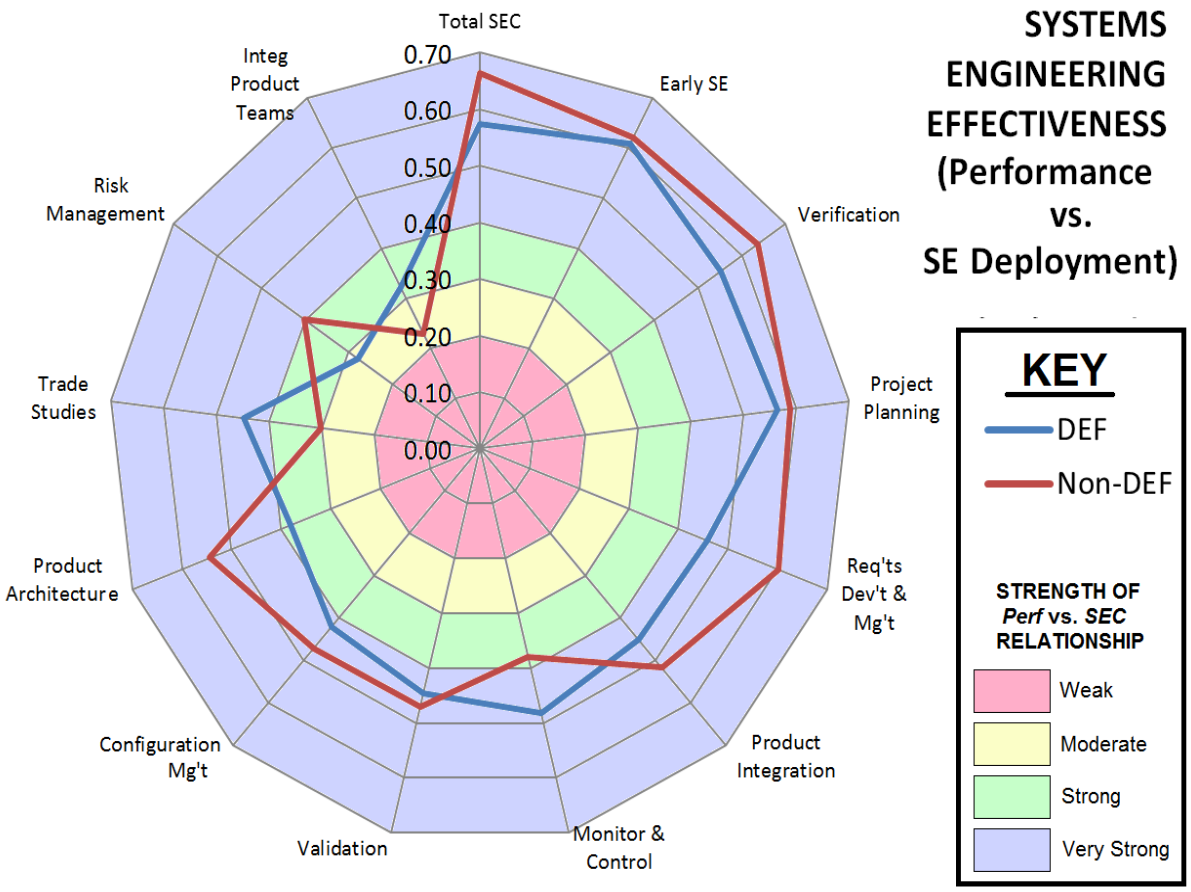
SYSTEMS ENGINEERING DEPLOYMENT



PROJECT PERFORMANCE



Defense vs. Non-defense Projects -2



Next Steps: Investigate the differences between SE deployment / effectiveness in defense and non-defense domains to find "transplantable" best practices

Next Steps

Download the 2012 SE Effectiveness reports from the SEI website

http://www.sei.cmu.edu/measurement/research/acquisition/Business-Case-SE.cfm

- The Business Case for Systems Engineering Study: Results of the Systems Engineering Effectiveness Survey
- The Business Case for Systems Engineering Study: Detailed Response Data
- The Business Case for Systems Engineering Study: Assessing Project Performance from Sparse Data
- The Business Case for Systems Engineering: Comparison of Defense Domain and Non-Defense Projects

Search for ways to apply the findings within your own work and your own organization.

Contact the SEI with questions or to obtain assistance.

SEI – Your Resource for Software and Systems Engineering

For more information, contact

Software Engineering Institute
4500 Fifth Avenue
Pittsburgh, PA 15213-2612

info@sei.cmu.edu
412-268-5800
1-888-201-4479

OR

Joseph P. Elm

jelm@sei.cmu.edu
412-268-9132

Software Engineering Institute | Carnegie Mellon University

NATIONAL DEFENSE INDUSTRIAL ASSOCIATION
NDIA
STRENGTH THROUGH INDUSTRY & TECHNOLOGY

AESS tm

BACK UP

References

Elm, J.; Goldenson, D.; El Emam, K.; Donatelli, N.; Neisa, A. "***A Survey of Systems Engineering Effectiveness – Initial Results***". Carnegie Mellon University; Pittsburgh, PA. 2007 (available at http://resources.sei.cmu.edu/library/asset-view.cfm?assetid=8493)

Elm, J.; Goldenson, D. "**The Business Case for Systems Engineering Study: Results of the Systems Engineering Effectiveness Survey**". Carnegie Mellon University; Pittsburgh, PA 2012 (available at http://resources.sei.cmu.edu/library/asset-view.cfm?assetid=34061)

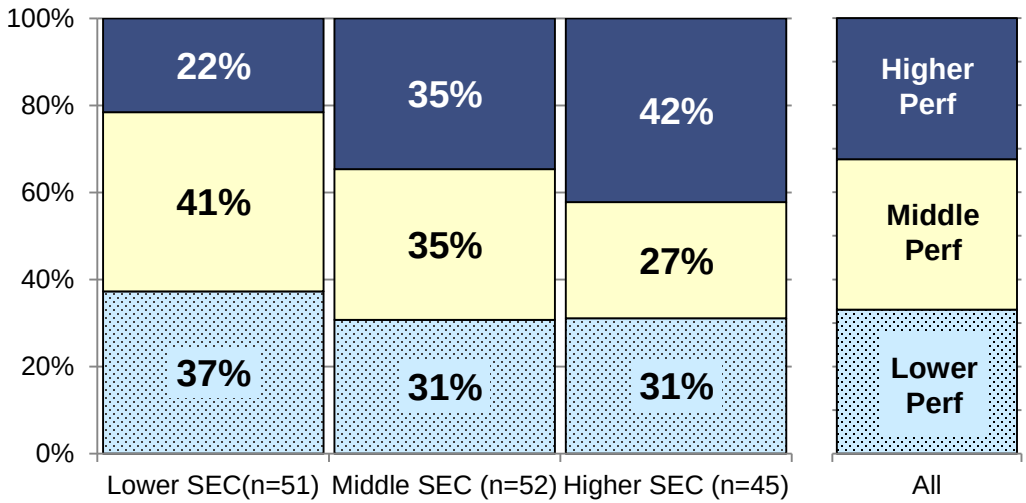
Elm, J.; Goldenson, D. "**The Business Case for Systems Engineering Study: Detailed Response Data**". Carnegie Mellon University; Pittsburgh, PA 2012 (available at http://resources.sei.cmu.edu/library/asset-view.cfm?assetid=73582)

Elm, J. "**The Business Case for Systems Engineering Study: Assessing Project Performance from Sparse Data**". Carnegie Mellon University; Pittsburgh, PA 2012 (available at http://resources.sei.cmu.edu/library/asset-view.cfm?assetid=34055)

Elm, J.; Goldenson, D. "**The Business Case for Systems Engineering: Comparison of Defense Domain and Non-Defense Projects**". Carnegie Mellon University; Pittsburgh, PA 2014 (available at http://resources.sei.cmu.edu/library/asset-view.cfm?assetid=?????)

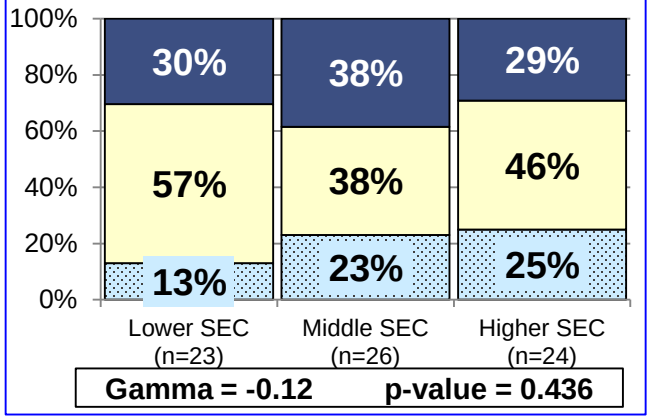
Integrated Product Teams vs. Performance

Perf vs. SEC-IPT



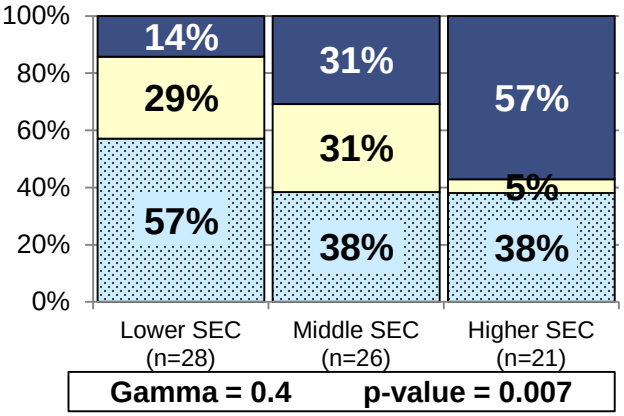
Gamma = 0.18 p-value = 0.101

Perf vs. SEC-IPT (Low PC)



Gamma = -0.12 p-value = 0.436

Perf vs. SEC-IPT (High PC)



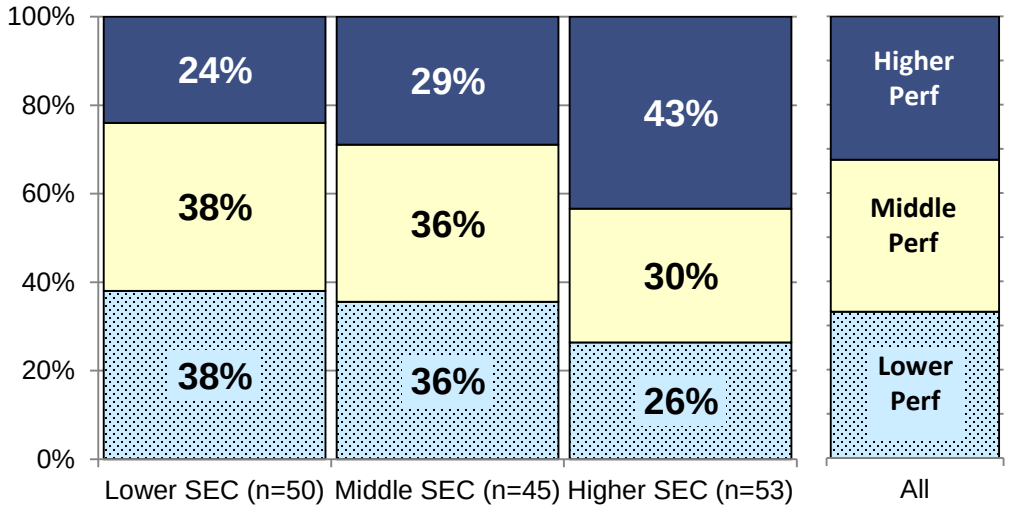
Gamma = 0.4 p-value = 0.007

The relationship:

for the set of all programs	**0.18 = Weak**
for the set of Low Challenge programs	**-0.12 = Weak Neg.**
for the set of High Challenge programs	**0.40 = Strong**

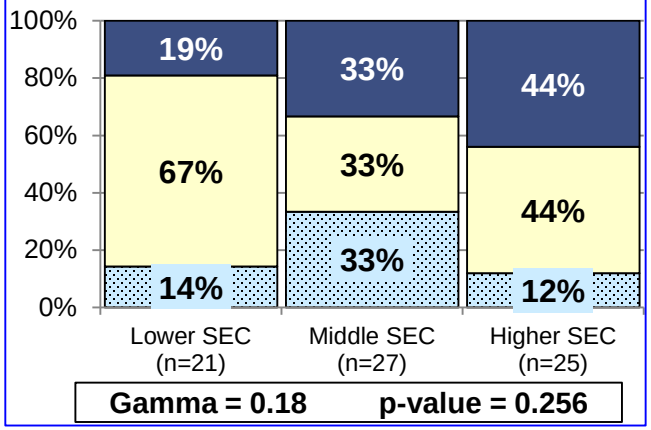
Risk Management vs. Performance

Perf vs. SEC-RSKM



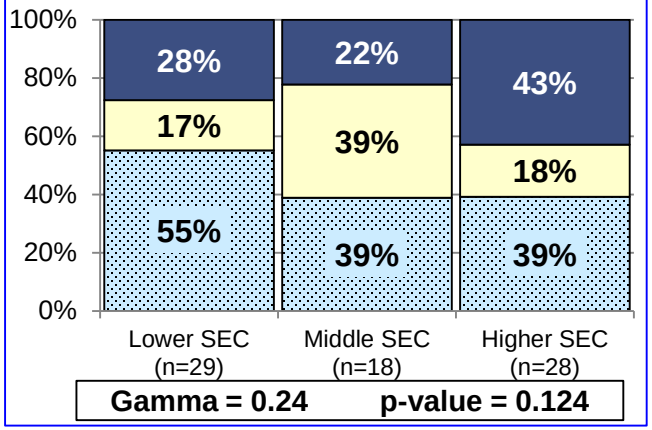
Gamma = 0.21 p-value = 0.05

Perf vs. SEC-RSKM (Low PC)



Gamma = 0.18 p-value = 0.256

Perf vs. SEC-RSKM (High PC)



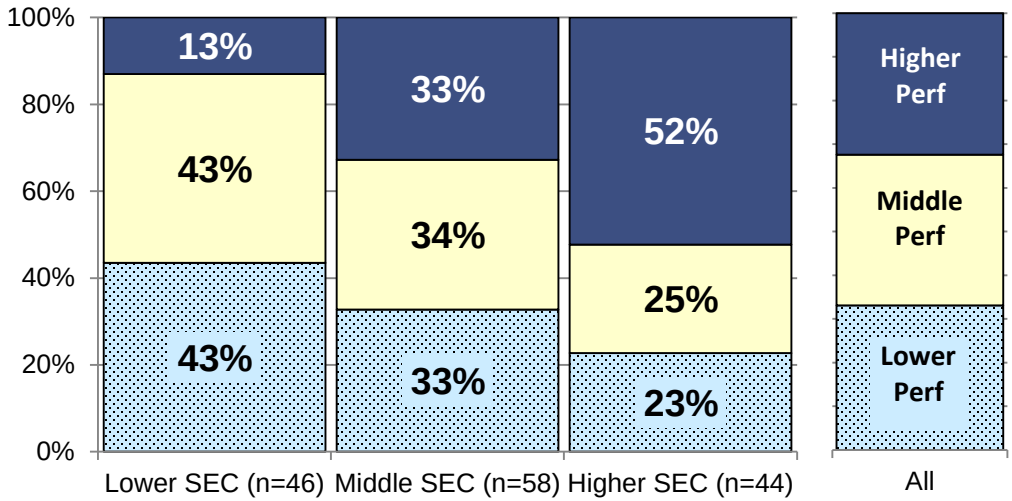
Gamma = 0.24 p-value = 0.124

The relationship:

for the set of all programs	**0.21 = Moderate**
for the set of Low Challenge programs	**0.18 = Weak**
for the set of High Challenge programs	**0.24 = Moderate**

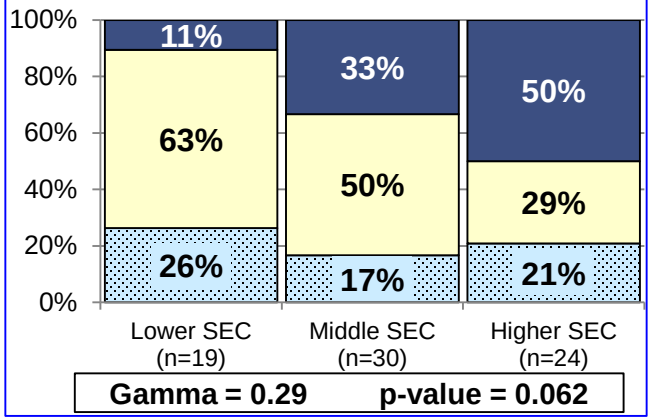
Trade Studies vs. Performance

Perf vs. SEC-TRD



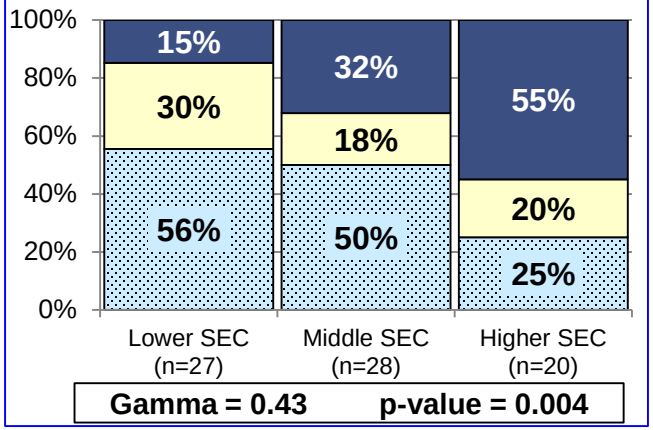
Gamma = 0.38 p-value = 0

Perf vs. SEC-TRD (Low PC)



Gamma = 0.29 p-value = 0.062

Perf vs. SEC-TRD (High PC)



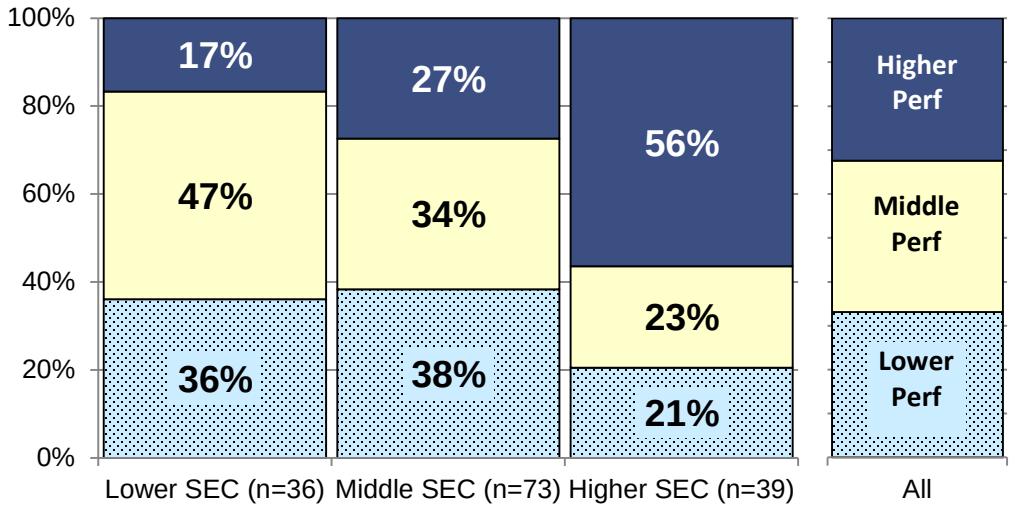
Gamma = 0.43 p-value = 0.004

The relationship:

for the set of all programs	**0.38 = Strong**
for the set of Low Challenge programs	**0.29 = Moderate**
for the set of High Challenge programs	**0.43 = Very Strong**

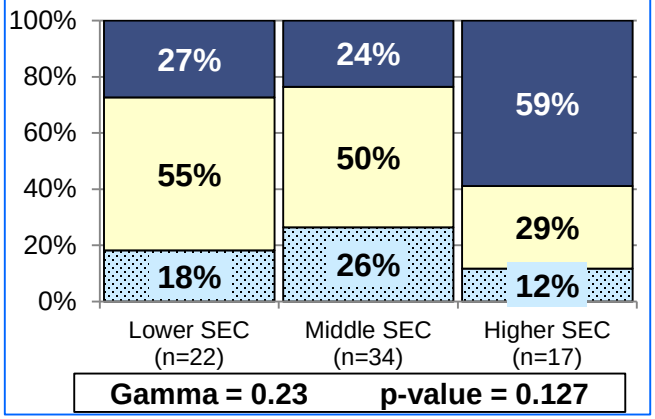
Validation vs. Performance

Perf vs. SEC-VAL



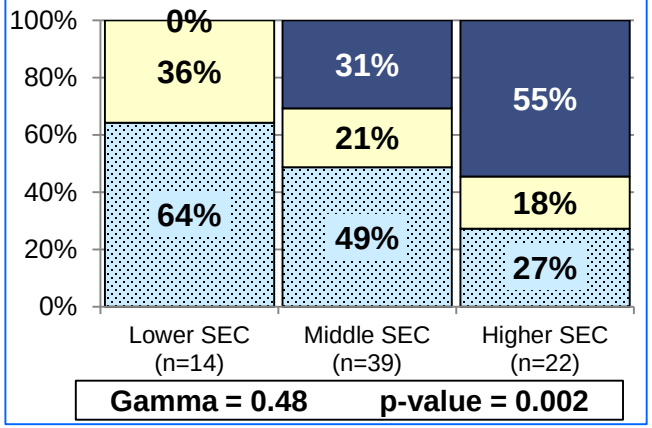
Gamma = 0.33 p-value = 0.003

Perf vs. SEC-VAL (Low PC)



Gamma = 0.23 p-value = 0.127

Perf vs. SEC-VAL (High PC)



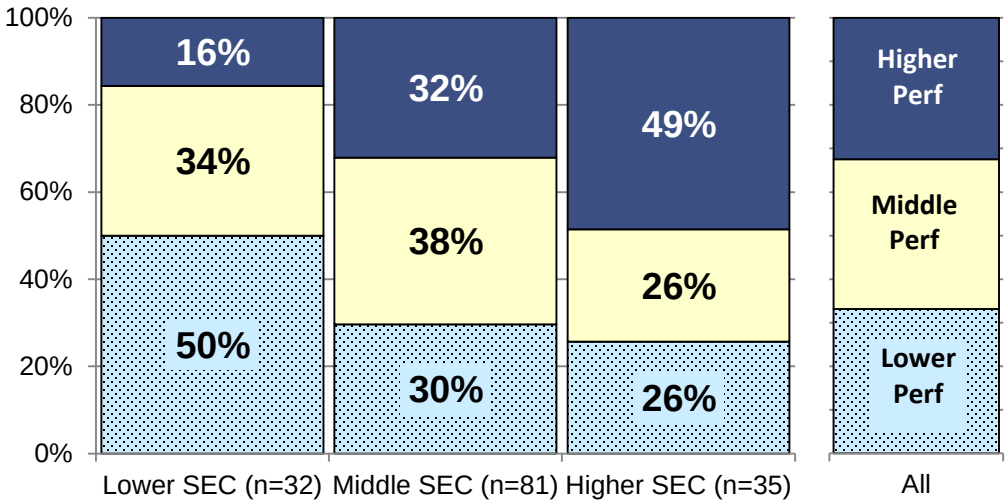
Gamma = 0.48 p-value = 0.002

The relationship:

for the set of all programs	**0.33 = Strong**
for the set of Low Challenge programs	**0.23 = Moderate**
for the set of High Challenge programs	**0.48 = Very Strong**

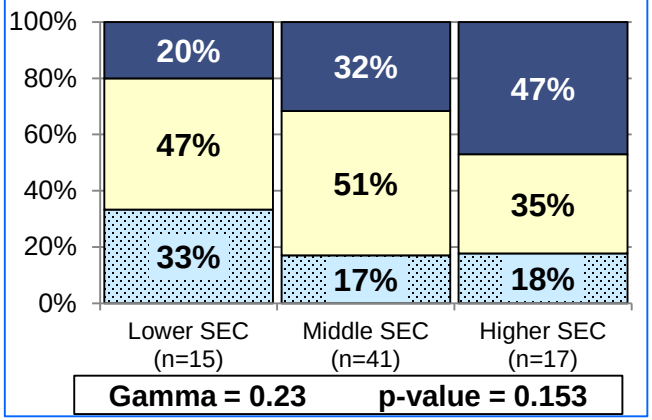
Product Integration vs. Performance

Perf vs. SEC-PI



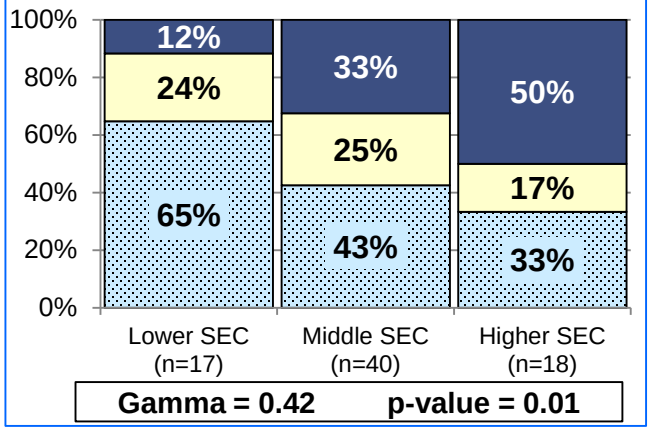
Gamma = 0.33 p-value = 0.003

Perf vs. SEC-PI (Low PC)



Gamma = 0.23 p-value = 0.153

Perf vs. SEC-PI (High PC)



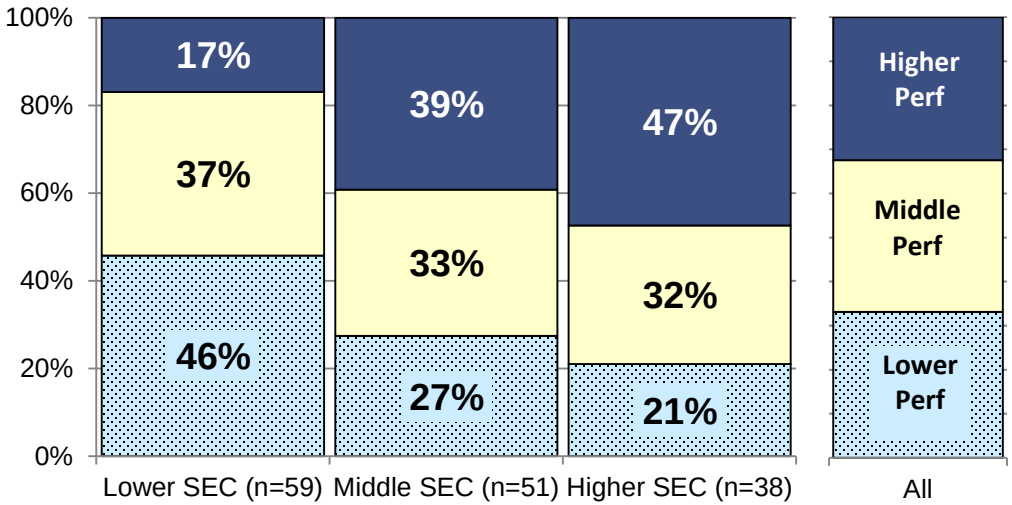
Gamma = 0.42 p-value = 0.01

The relationship:

for the set of all programs	**0.33 = Strong**
for the set of Low Challenge programs	**0.23 = Moderate**
for the set of High Challenge programs	**0.42 = Very Strong**

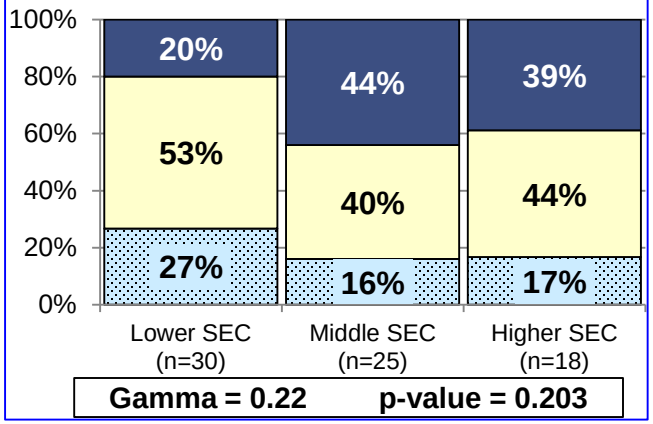
Configuration Management vs. Performance

Perf vs. SEC-CM



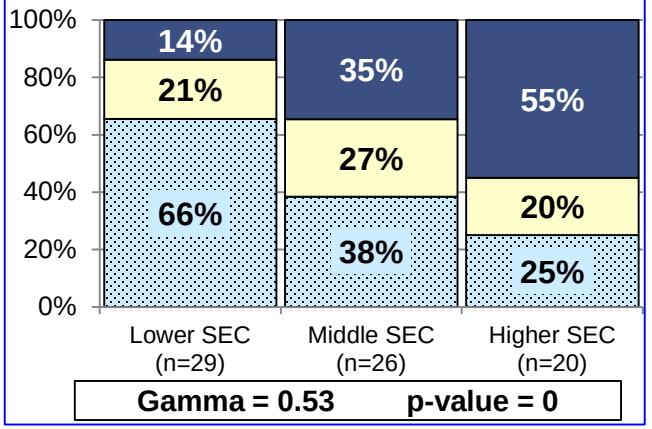
Gamma = 0.38 p-value = 0.001

Perf vs. SEC-CM (Low PC)



Gamma = 0.22 p-value = 0.203

Perf vs. SEC-CM (High PC)



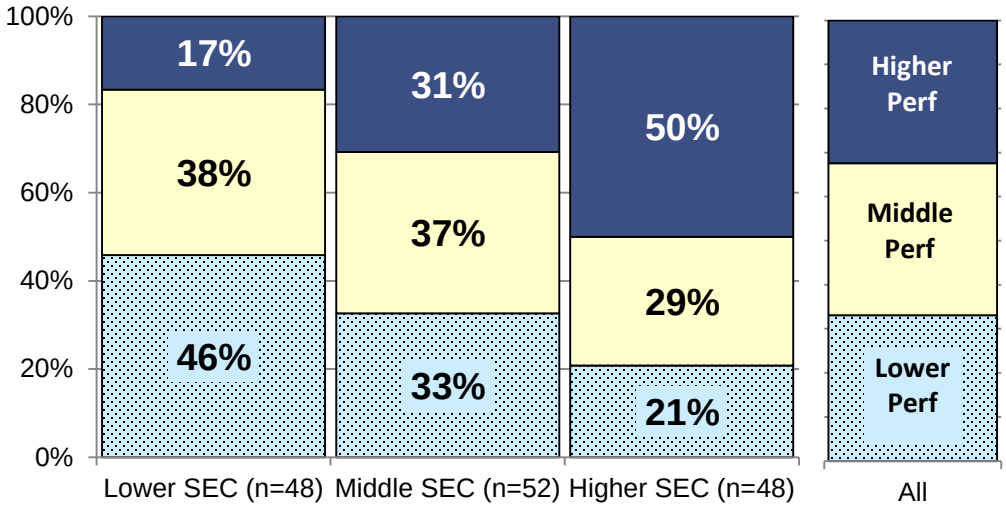
Gamma = 0.53 p-value = 0

The relationship:

for the set of all programs	**0.38 = Strong**
for the set of Low Challenge programs	**0.22 = Moderate**
for the set of High Challenge programs	**0.53 = Very Strong**

Program Monitoring & Control vs. Performance

Perf vs. SEC-PMC

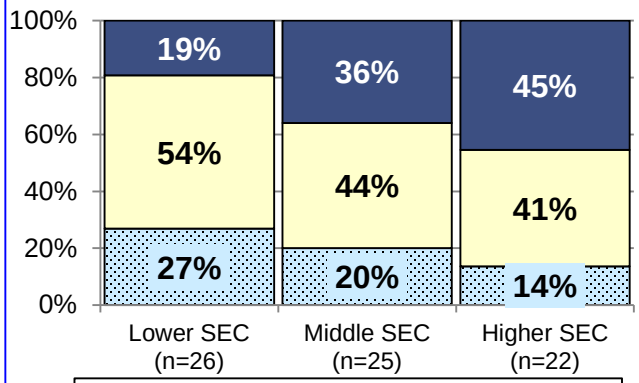


Gamma = 0.38 p-value = 0

The relationship:

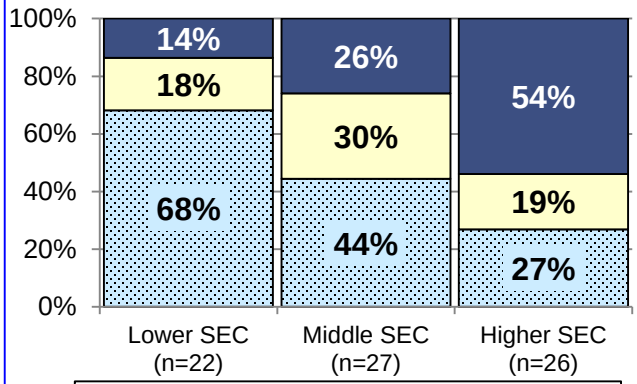
for the set of all programs	**0.38 = Strong**
for the set of Low Challenge programs	**0.27 = Moderate**
for the set of High Challenge programs	**0.53 = Very Strong**

Perf vs. SEC-PMC (Low PC)



Gamma = 0.27 p-value = 0.092

Perf vs. SEC-PMC (High PC)



Gamma = 0.53 p-value = 0