



26th Annual ASQ Audit Conference
Blueprint for a Successful Audit

October 9 - 11, 2017 • Pre-conference Courses
October 12 - 13, 2017 • Conference

ASQ

The Global Voice of Quality™ The InterContinental Hotel | Dallas, Texas



ASQ®

Audit
Division

The Global Voice of Quality™

Sampling for Software Process Assessments, Evaluations, and Appraisals

Dr. Mark C. Paulk
13 October 2017

What Is An Assessment?

- An appraisal of an organization's current software process for self-improvement by a trained team of experienced software professionals.
- It is based on
 - review of 4 to 6 representative projects
 - responses to the maturity questionnaire
 - in-depth discussions with project managers and practitioners
 - collective knowledge and experience of the assessment team

The “Evolution” of CMMs

- Software process maturity framework 1987
- Software CMM v1.1 1993
 - Software Acquisition CMM
 - Systems Engineering CMM
 - ... over 60 different CMMs – staged and continuous
- eSCM for Service Providers v1 2001
- CMMI for Development v1.1 2002
 - CMMI for Acquisition
 - CMMI for Services
- eSCM for Service Providers v2 2004
- CMMI for Development v1.3 2010

The “Evolution” of Assessments (1 of 2)

- Software process assessments (SPA)
 - software process improvement, 1987 to 1995
 - “What are your problems?”
 - map problems to the maturity framework
 - emphasis on organizational intervention
- Software capability evaluations (SCE)
 - source selection, contract monitoring
 - audits against the maturity framework, 1988 to 1995
 - audits against the Software CMM, 1995 to 2003
 - emphasis on level playing field – fairness
- Appraisals = assessments and evaluations

The “Evolution” of Assessments (2 of 2)

- CMM-based appraisals for internal process improvement (CBA IPI)
 - software process improvement, 1995 to 2003
 - audit process against Software CMM
 - identify problems, map against Software CMM as appropriate
 - report CMM and non-CMM findings
 - emphasis on reliable and consistent results
- Standard CMMI appraisal method for process improvement (SCAMPI)
 - process improvement, 2003 till ...
 - audits against CMMI

Assessment Versus Evaluation

Issue	Assessment	Evaluation
Use	process improvement	source selection
Objective	assess current practice	substantiate practice
Improvement goal	catalyst for improvement	evaluate commitment
Output	input for action plan	performance risk
Range of findings	non-CMM findings possible	CMM findings only
Style	collaborative	audit-oriented
Focus of results	applies to organization	predict next project
Status of results	confidential	known to DoD

SEI's SCAMPI v1.3

- “Standard CMMI Appraisal Method for Process Improvement (SCAMPI) A, Version 1.3: Method Definition Document,” Carnegie Mellon University, Software Engineering Institute, CMU/SEI-2011-HB-001, March 2011.
- “Sampling is planned quantitatively based on
 - the diversity of unique process implementations within the appraisal scope,
 - with the goal of both ensuring a representative sample of the organizational unit and
 - optimizing the effort for collection and analysis of objective evidence.”

Organizational Analysis

- The scope of what “organization” means can cover
 - Multiple business units, product lines, domains (type of work)
 - Customer segmentation
 - Sizes of project and/or product, duration
 - Software engineering methods and environments
 - Geographical sites (location)
 - Workforce mixes (experience, ...)
 - Organization vs project-level processes

SCAMPI Sampling Units

- Basic units ~ projects + support functions (organizational functions), ...
- Organizational scope of the appraisal is selected as a representative sample of the organizational unit.
 - based on sampling factors that reflect meaningful differences in the conditions under which work is performed

SCAMPI Sampling Example

	Commercial customers	Government customers
New York	7	2
Cincinnati	5	0
Denver	11	5

	Total number of basic units	Basic units sampled
New York, Commercial	7	1
New York, Government	2	1
Cincinnati, Commercial	5	1
Denver, Commercial	11	2
Denver, Government	5	1

The Problem

- Is this an adequate sampling approach?
- Performing assessments in large complex organizations quickly becomes cost prohibitive if every component (entity) must be assessed for each criterion
- Sampling methodology goals
 - Reduce the number of entities that must be assessed, especially for assessments with a large number of entities
 - Keep the risk of incorrectly certifying an organization that is not really compliant at or below 10% under most circumstances

Sampling

- Sampling allows one to specify the amount of error one can tolerate, thus maintaining the desired accuracy levels while substantially reducing the time and effort involved.
- “Four to six representative projects” might be considered a “convenience sample.”
- Covering all aspects of the organization and its projects is a distinct improvement but does not address the question of adequate coverage.

Statistical Sampling Prerequisites

- *Randomization*, every sampling unit (project) must have the same likelihood of being selected as part of any given sample.
- *Stratified random sample*, a sample obtained by separating the population elements into non-overlapping groups (strata), and then selecting a simple random sample from each stratum.

No Guarantees Assumption

- The ability within a stratum to detect a single entity that is not compliant was not a design goal.
 - A high likelihood of finding at least one noncompliant entity when at least two exist within a stratum.
 - It does not guarantee finding a single non-compliant entity.
 - It does not guarantee finding multiple non-compliant entities.

Perfect Team Assumption

- Any entity that is truly not compliant and is selected in the sample will be detected.

The Ultimate Sampling Objective

- Incorrectly passing a single Practice, while undesirable, is not as serious as incorrectly passing an organization at any given Capability Level.

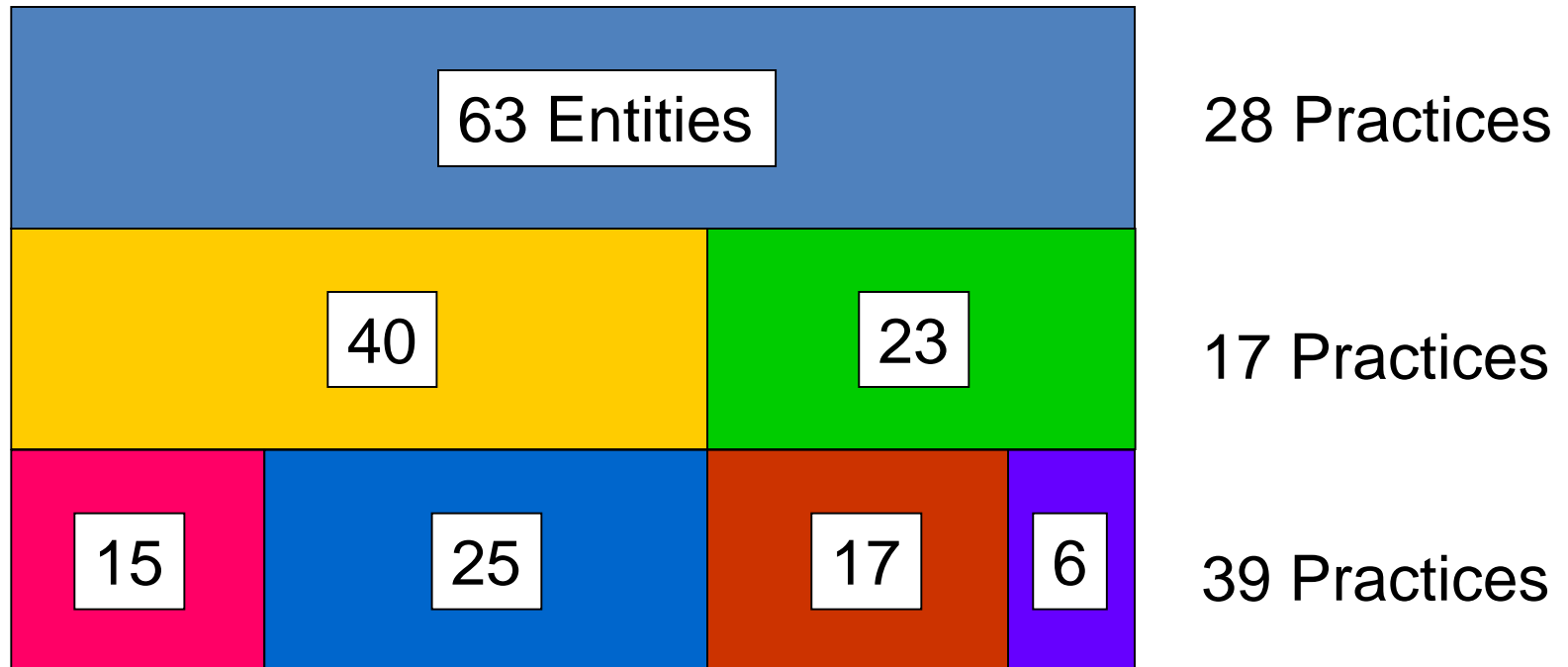
Sampling Method

- Uses a variation of Stratified Random Sampling – entities are placed in similar strata by Practice (84 Practices in eSCM-SP)
- Stratifying at the Practice level allows us to differentiate while still achieving maximum efficiency for Practices that are implemented across large parts of the organization (e.g., HR, Contracting, etc.)

Stratifying the Population

- Identify the entities that are in scope for the assessment.
- Organize entities into strata that have similar characteristics (organizational analysis).
 - Consistent process implementations within a strata (homogeneous subgroups)
 - Responsibility of Lead Assessor
 - Approved by certification body

Example Stratification for 63 Entities



Sampling Table Design

- Designed to control the probability of an organization incorrectly achieving a Capability Level for which they were not truly compliant.
- Set the sample sizes such that the conditional probability is less than 10% that no Practice noncompliances will be detected, given that Practice noncompliances exist.
- If any sample would have a 70% or better chance of detecting at least one noncompliant entity, when there are two or more noncompliant entities in the stratum, we will achieve a minimum of a 91% probability of detecting at least one Practice noncompliance.

eSCM Sampling

If any sample would have a 70% or better chance of detecting at least one noncompliant entity, when there are two or more noncompliant entities in the stratum, we will achieve a minimum of a 91% probability of detecting at least one Practice noncompliance.

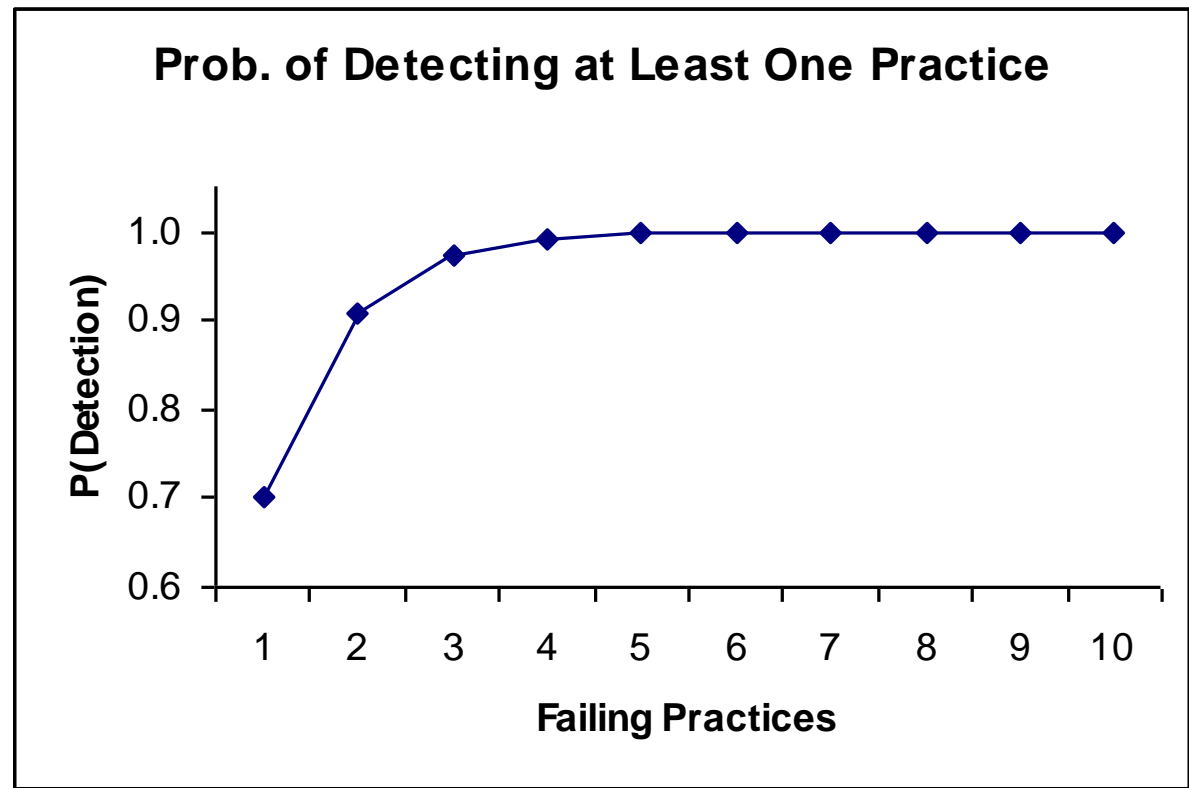
Number of Entities	Sample Size	Number of Entities	Sample Size	Number of Entities	Sample Size
1-3	All	21-22	10	39-40	18
4-7	3	23-24	11	41-42	19
8-9	4	25-27	12	43-44	20
10-11	5	28-29	13	45-46	21
12-13	6	30-31	14	47-49	22
14-16	7	32-33	15	50 and up	23
17-18	8	34-35	16		
19-20	9	36-38	17		

A Simple eSCM Sampling Example

Stratum	Number of Entities in Stratum	Number of Practices	Sample Size
Org	93	17	23
S1	43	67	20
S2	25	67	12
S3	17	67	8
S4	8	67	4

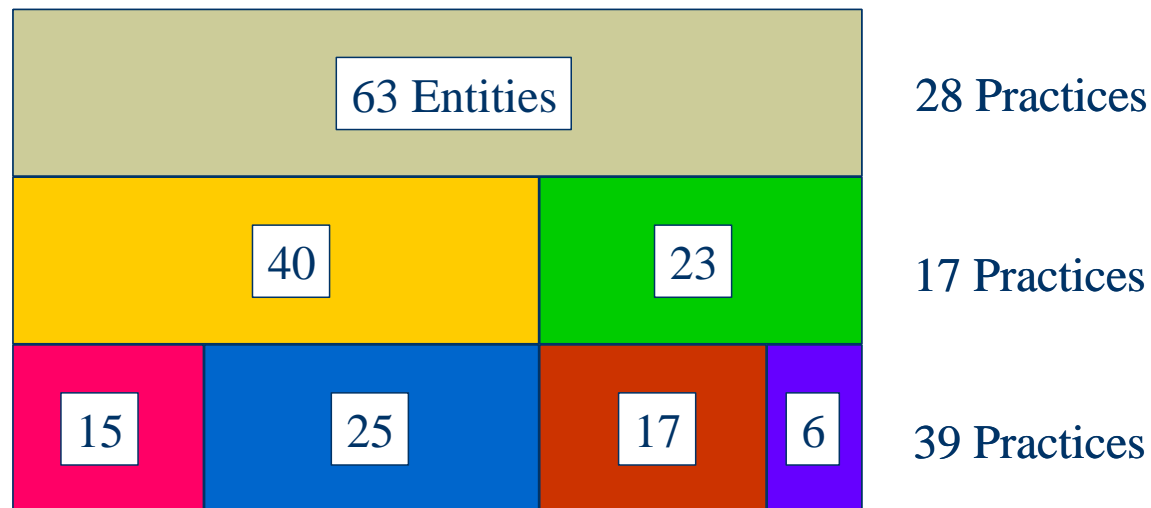
Failed Level Detection

Because a single failed practice is sufficient to fail at any Level, detecting Level failure is the same as detecting Practice failure.



Typical Results

- With a full census, the earlier example would have required 5,292 Entity-Practice pairs to be reviewed
- With the sampling plan, only 2,307 need to be reviewed – a 56.4% reduction



Conclusions

- It is possible to significantly reduce the effort required to evaluate organizations with multiple entities while still controlling for the number of false positives
- While the reduction is significant in relative terms, the remaining effort for large organizations is still substantial in absolute terms

eSCM

- E.B. Hyder, K.M. Heston, M.C. Paulk, and W.E. Hefley, The eSourcing Capability Model for Service Providers, Van Haren Publishing, Zaltbommel, Netherlands, 2009.
- D.M. Northcutt and M.C. Paulk, “Statistical Sampling for Process Assessments,” ASQ Software Quality Professional, 12 (4), September 2010, pp. 19-28.
 - “By sampling we mean the process of drawing an inference about some population of interest by examining only a subset of the population.”

Questions?