#### Software Solutions Symposium 2017 March 20–23, 2017

#### Best Practices for Software Process and Product Measurement

#### **Dr. Bill Curtis**

Executive Director, CISQ SVP & Chief Scientist, CAST Research Labs

Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213



# CISQ Instructor — Dr. Bill Curtis



Dr. Bill Curtis is Senior Vice President Chief Scientist of CAST Software, where he heads CAST Research Labs. CAST is a leading provider of technology for measuring the structural quality of business applications. He leads the production of CAST's annual CRASH Report on the global state of structural quality in business applications. He is also the Director of the Consortium for IT Software Quality, a Special Interest Group of the Object Management Group (OMG) where he leads an international effort to establish standards for automating measures of software quality characteristics. Dr. Curtis is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE) for his career contributions to software process improvement and measurement.

Dr. Curtis is a co-author of the Capability Maturity Model for Software (CMM), the People CMM, and the Business Process Maturity Model. He was also a member of the Product Suite Team for CMMI. He is a globally recognized expert in software measurement, process improvement, and workforce development. He has been leader in the application of statistics to software engineering. He was the Chief Process Officer at Borland Software Corporation. Until it's acquisition by Borland he was a Co-founder and Chief Scientist of TeraQuest in Austin, Texas, the premier provider of CMM-based services. He is a former Director of the Software Process Program in the Software Engineering Institute at Carnegie Mellon University. Prior to joining the SEI, Dr. Curtis directed research on advanced user interface technologies and the software design process at MCC, the US's fifth generation computer research center in Austin, Texas. Prior to MCC he developed a global software productivity and quality measurement system at ITT's Programming Technology Center. Prior to ITT he evaluated software development methods in GE Space Division, worked in organizational effectiveness at Weyerhaeuser, and taught statistics at the University of Washington. He has co-authored four books, is on the editorial boards of several journals, and has published over 150 papers on software development and management. Dr. Curtis earned a Ph.D. from Texas Christian University and a M.A. from The University of Texas. He earned his B.A. from Eckerd College in mathematics, psychology, and theater.



section	slide
Measuring Software Size and Productivity Break	3
Adopting Measures to Development Methods Lunch	33
Measuring Software Improvement Programs Break	61
Measuring Software Product Quality Adjourn	89
References	115
	section Measuring Software Size and Productivity Break Adopting Measures to Development Methods Lunch Measuring Software Improvement Programs Break Measuring Software Product Quality Adjourn References

- © 2017. The presentation material in these tutorial notes is copyrighted by Dr. Bill Curtis. It is to be used for the benefit of the individuals attending this tutorial. It may not be copied or distributed to others without written permission from Dr. Curtis. For further information, please contact:
- Capability Maturity Model, Capability Maturity Model Integration, People Capability Maturity Model, CMM, and CMMI are registered in the U.S. Patent and Trademark Office
- Dr. Bill Curtis P.O. Box 126079 Fort Worth, Texas 76126-0079 USA <u>curtis@acm.org</u>



# Measuring Software Size and Productivity

- **1. Practical Software Measurement**
- 2. Size and functional measures
- 3. Measuring software productivity
- 4. Analyzing software productivity

# CISQ Why Does Measurement Languish?





Too many measures



Use in personal appraisals





**Unreliable data** 

#### Measurement Process Guidance

- Based on decades of software measurement best practice
- Best guidance for starting a measurement program
- Compatible with *ISO 15939 Software Measurement Process*
- Free guides & licensed training and consulting available
- http://psmsc.com/



McGarry, et al., (2002). Practical Software Measurement.

# Practical Software Measurement

Objective Information for Decision Makers

> John McGarry David Card Cheryl Jones Beth Layman Elizabeth Clark Joseph Dean Fred Hall

# CISQ Measurement Categories

	Category	lssue	Example measure	
vel 2	Schedule & Progress	Milestone completion Earned value	% Actual vs. planned dates Actual vs. planned completions	
	Resources & Cost	Personnel effort Facilities used	Person-hours Test equipment & hours	
	Product Size	Size of coded product Size of documentation	Function Points Pages of manuals	
3 Le	Product Quality	Functional quality Structural quality	Defects per KLOC Violations of reliability rules	
≥ Level	Process Performance	Efficiency Compliance	Defect detection efficiency Audit results	
	Product Value	Return on investment Risk avoidance	$\Delta$ Increase in target revenue Cost of system outage	
	Customer Satisfaction	Account growth Self–sufficiency	Repeat business within 1 year Number of helpdesk calls	

McGarry, et al., (2002). Practical Software Measurement, p.37

# CISQ Productivity Analysis Objectives



# CISQ Productivity Analysis Measures





**Productivity** 



Total effort expended on the release



#### **Software Size Measures**

#### Instructions

#### Lines of Code

Most frequently used. Different definitions of a line can cause counts to vary by 10x. Smaller programs often accomplish the same functionality with higher quality coding.

<b>Requirements-based</b>	Use Case Points, Story Points
---------------------------	-------------------------------

Use Case Points have not become widely used and need more development. Story points are subjective to each team and are susceptible to several forms of bias.

#### **Functions**

#### **Function Points**

Popular in IT. Several counting schemes (IFPUG, NESMA, Mark II, COSMIC, etc.). Manual counting is expensive and subjective—certified counters can differ by 10%. Automated FPs taking root.

#### How Many Lines of Code ? Consortium for IT Software Quality



B. Park (1992)

# **CISQ** Function Point Estimation





Functional view of software

- Functional size can be estimated from external inputs and outputs
- Upfront functional analysis provides basis for good estimates
- Repository of FP data provides basis for calibrating estimates

## CISQ Effort — Weakest Measure

# After the fact estimates

- Memory lapses
- Time-splicing
- Inconsistency

Underreporting

- Contract issues
- HR issues
- Impressions

# Lack of normalization

- Roles included
- Phases included
- Hours in P-Year

Effort Unreliable, Inconsistent

# CISQ How Quality Affects Productivity

#### **Assumption**: Productivity is a stable number

**Reality**: Productivity is unstable, tending to decline



#### Unless you take action !!!

### CISQ Carry-forward Rework



#### **Example of Quality Impact**

#### Project A (Plodders)

- 20 developers, 3 months
- \$120k per FTE
- 3 FPs per staff month
- 180 FPs delivered
- \$3,333/FP cost

#### Project B (Better, Faster, Cheaper)

- 20 developers, 3 months
- \$120k per FTE
- 4 FPs per staff month
- 240 FPs delivered
- \$2,500/FP cost

#### Project B is 25% more productive

#### However !!!

- 2 critical violations per FP
- \$500 per fix
- Cost for 360 fixes = \$180k
- Total Cost to Own = \$780k
- \$4,333/FP of TCO

- 5 critical violations per FP
- \$500 per fix
- Cost for 1200 fixes = \$600k
- Total Cost to Own = \$1,200k
- \$5,000/FP of TCO

#### Project A is 13.4% more productive

#### **Quality-Adjusted Productivity**

Consortium for IT Software Quality



#### **Best Practices for Analysis**

- 1) Segment baselines
- **Beware sampling effects** 2) **Understand** variation **Evaluate demographics** 4) 5) Investigate distributions Account for maturity effects 6) Beware external data sets 7)

# CISQ 1 — Evaluate Demographics





#### **Function Points**

# CISQ 2 — Segment Baselines

#### Multiple baselines are usually the most valid

Year	Projects	Productivity
Total Corporate		
1981	28	2342
1980	21	1939

# CISQ 3 — Beware Sampling Effects



#### 4 — Understand Variation Consortium for IT Software Oualit



- complexity
- product size

J. Vosburgh, B. Curtis, et al. (1984). Proceedings of ICSE 1994.

.65

# **CISQ** 5—Investigate Distributions



# CISQ 6 — Account for Maturity Effects



Which organization will be required to spend more effort on correcting software flaws?

# CISQ 7 — Caution for External Data Sets



- ? Data definitions
- ? Data collection
- Data validity
- ? Data age
- ? Demographics

# Section 2 Measuring to Manage Software Projects

- 1) Tracking traditional projects
- 2) Iterative and Agile measurement
- 3) Team Software Process measurement

#### CISQ Measuring Project Progress







#### CISQ Phase Defect Removal Model

PHASE	Escapes Previous Phase / KSLOC	Defect Injection / KSLOC	Subtotal / KSLOC	Removal Effective- ness	Defect Removal / KSLOC	Escapes at phase exit / KSLOC
Rqmts	0	1.2	1.2	0%	0	1.2
Design	1.2	18.0	19.2	76%	14.6	4.6
Code & U.T.	4.6	15.4	20.0	71%	14.2	5.8
Inte- gration	5.8	0	5.8	67%	3.9	1.9
Sys. Test	1.9	0	1.9	58%	1.1	0.8
Opera- tion	0.8					

Kan (1999). Metrics and Model in Software Quality Engineering.

# CISQ SCRUM Board

III Virtual SCRUM Board							
Elle Edit View Sprint Story Task Tools Help							
🤹 Planning 👹 Current Spr	🖼 Planning 🕼 Current Sprint 🕮 Scrum Board 🕮 Sprint Burndown 🗠 Print Index Cards						
	3/3/2008 - 3/14/2008						
Story	Not Started	In Progress	To Verify	Done			
As a CFO, I would like to see a monthly sales trend. [JPN]	6 Test report view and printing 3	Impeded 8 Modify UI to allow selection of new report		8 Understand the facilities of the existing query engine. [JPN]			
	Update application documentation and rebuild help file			4 Implement new report template [EB]			
5 As an Accountant, I would like to create invoices for all	4 Write Unit Tests to check data selection criteria	12 Write code to select appropriate data for invoicing	4 Design Invoice template	0 Research reporting engines			
outstanding charges for all customers	8 Implement UI	2 Wireframe UI for the Invoice Run		0 Review the viable choices for Reporting Engines with the team for input			
	8 Implement and run Functional Tests 6			0 Learn API of the reporting engine			
C:\Dev\Code\VirtualScrumBoard\VSBDemoData\VSBData 7 Valued Customer (Single-User License)							

- Scrumboard displays the path to completion for each story by showing task status
- Story points are an estimate of days for each story
- Agile estimating is by feature (story) rather than by task
- 'Done' does not always mean all tasks are finished
- Kanban methods restrict the work in progress to a limited number of paths

# CISQ Burndown Chart and Velocity



- Burndown chart displays story completion by day
- Burndown chart indicates actual versus estimated progress
- Velocity is the rate at which stories are completed
- Velocity indicates sustainable progress
- Velocity results provide one source of historical data for improving estimates
- Story points without the context of productivity factors are dangerous for estimating

#### **Analyzing Velocity** Consortium for IT Software Quality

#### Burndown Chart



# CISQ Measuring and Managing Flow





Anderson (2010). Kanban.

# CISQ Trends Over Sprints





#### **Data sources**

Project tracking

Code control

**Bug tracking** 

**Build system** 

Test environment

**Deployment tools** 

**Operational monitoring** 

#### Measures

Stories submitted

**Stories pulled back** 

Growth in size of stories

Stories added mid-sprint

Stories under review

**Failed tests** 

Non-development tasks

Assists to other projects


#### **Recommended Books**



#### Personal Software Process (PSP)



Humphrey (1997). Introduction to the Personal Software Process.

Consortium for IT Software Quality

**Team Software Process (TSP)** 

- Built from personal processes of team members
- Well defined team roles
- Project launch workshops for planning
- Team measurement and tracking
- Post-mortems for improvement
- Team application of lean principles



Humphrey (2000). Introduction to the Team Software Process.

#### **Aggregated Personal Data Best**



Consortium for IT Software Quality

The ability to predict the amount of time required to produce a piece of software is dramatically improved by estimating at the individual level first and then combining the estimates rather than developing a single team estimate averaged over individuals.

#### CISQ Defect Detection at Intuit



Fagan (2005)



#### 2.0 \* **Dramatic reductions:** 1.5 Delivered defects defects/KLOC Days per KLOC 1.0 - Schedule deviation Test defects/kloc 0.5 Variation in results 0 0.0 N = 9 12

Without TSP

With TSP

# Section 3 Measuring Software Improvement Programs

1) Improvement program methods

- 2) Maturity-based measurement
- 3) Improvement program results

### CISQ CMMI Maturity Transitions



Chrisis, Konrad, & Shrum (2005). CMMI.

#### **Enhanced Maturity Framework** Consortium for IT Software Quality

#### Disputation of Misinterpreted Principles Underlying the Process Maturity Framework

August 1, 2011

Out of respect for the Process Maturity Framework the following propositions are presented to diapate the mininterpretation of its founding principles and to eliminate the confusion that currently afflicts maturity models and the practice of process improvement. The following propositions were discussed at Dubin during the keynote of Dr. Bill Curris. Wherefore he requests that those who were mubble to be present and debate collizions of on electromically.

In the name of Watts by whom enterprises are made mature. Amen

#### Theoretical Foundation

- The birth of Humphrey's Process Maturity Framework was in rejecting the practice-centric approach to process improvement first formulated in Crosby's Quality Maturity Grid and currently emoloxied in the continuous representation of ISO/IEC 15504 (SPICE) and Maturity Models based on this representation.
- Practice-centric approaches focus on improving individual processes or practices, often in isolation of each other, and they therefore fail to model an enterprise as an end-to-end system of interdependent processes.
- Decades of failed improvement programs have undermined the credibility of practice-centric approaches because weaknesses not addressed among interdependent processes eventually degraded the capability of processes improved in isolation of their dependencies.
- The capability and performance of an enterprise that functions as a collection of independent processes at different levels of maturity is usually determined by the capability of the least mature process, since it constrains the end-to-end system.
- 5. In rejecting the practice-centric approach and formulating the Process Maturity Framework as an enterprise-centric model currently embodied in the staged versions of Maturity Models, Humphrey established a new and original model of organizational development that has proven uniquely successful an guiding enterprise amprovement.
- The primary objective of an enterprise-centric Maturity Model is the continual and sustainable improvement of enterprise capability, where the term 'enterprise' represents the organizational infrastructure and resources required to perform the end-to-end value chain and support processes that produce and deliver the organization's products and services.



#### **CISQ**Maturity Level Transformation





Measure-maturity mismatch slows improvement and creates dysfunction

#### **Average Improvement Results**

Improvement benefit	Orgs	Annual median	Annual range
Productivity growth	4	35%	9% - 67%
Pre-test defect detection	3	22%	6% - 25%
Time to market	2	<b>↓ 19%</b>	15% - 23%
Field error reports	5	<b>↓ 39%</b>	10% - 94%
Return on investment	5	5.0:1	4:1 - 8.8:1

Consortium for IT Software Quality

#### CISQ Schedule Performance



### CISQ Cost Performance



### CISQ Crosby's Cost of Quality Model



#### CISQ Raytheon's Cost of Quality

- Performance *cost of building it right first time*
- Nonconformance *cost of rework*
- Appraisal cost of testing
- Prevention *cost of preventing nonconformance*

Year	Level	Perform	Nonconf.	Appraise	Prevent
1988	1	34%	41%	15%	7%
1990	2	55%	18%	15%	12%
1992	3	66%	11%	<b>§</b> 23	3% <b>\</b>
1994	4	76%	6%	<b>L</b> 18	s% <b>\$</b>

#### **Raytheon's Productivity**



Consortium for IT Software Quality

#### **Raytheon's Cost Reduction**



Consortium for IT Software Quality

### CISQ Raytheon's Cost Predictability



### CISQ OMRON's Reuse gains





# Measuring Quality and Technical Debt

- 1) Structural quality measurement
- 2) Technical debt
- 3) Consortium for IT Software Quality (CISQ)

#### **CISQ** What about the Product?

#### **CMMI's primary assumption:**

"The quality of a software system is governed by quality of the process used to develop it."

- Watts Humphrey

#### CMMI's product quality problem:

- 1. Assessments do not verify product quality
- 2. Compliance focus undermines CMMI's primary assumption
- 3. Product quality focus at Level 4 in CMM was lost in quantitative process control in CMMI
- 4. The assumptions underlying control charts are violated by software data

### CISQ Six Sigma's Challenge

#### Six Sigma projects must have significant benefits



**Process focus** – process improvement – *Six Sigma* **Product focus** – product improvement – *Design for 6* $\sigma$ 

Product focus supplements product focus to unlock even more value from software

### CISQ Testing is Not Enough





"As higher levels of assurance are
demanded...testing cannot deliver the level of confidence required at a reasonable cost."

"The correctness of the code is rarely the weakest link."

"...a failure to satisfy a non-functional requirement can be critical, even catastrophic...non-functional requirements are sometimes difficult to verify. We cannot write a test case to verify a system's reliability...The ability to associate code to non-functional properties can be a powerful weapon in a software engineer's arsenal."

#### **CAST's Application Intelligence Platform**

Consortium for IT Software Quality



#### Modern Apps are a Technology Stack



Consortium for IT Software Quality

## CISQ Architecturally Complex Defects



Most architecturally complex defects involve an architectural hotspot – a badly-built component that should be replaced

### CISQ Propagated Risk



#### **Transaction Risk**

#### Transaction Risk Index

Consortium for IT Software O

A ranking of transactions based on the risk created by the number and severity of violations along the transaction path

- Customers care most about the dependability of their transactions
- The risk of a transaction is determined by the number and severity of violations along its path
- Transactions can be ranked by risk to prioritize their remediation
- The Transaction Risk Index can be further tuned using the operational frequency of each transaction



Transaction 1 poses greatest risk



Violation	Health Factor	Severity	Propagated Risk Index	In a high risk transaction?	Frequency of path execution
#148	Security	9			
#371	Robustness	9			
#062	Performance	8			
#284	Robustness	7			
#016	Changeability	7			
#241	Security	5			
#173	Transferability	4			
#359	Transferability	3			





John Keane (2013).

#### Consortium for IT Software Quality Reducing Operational Incidents & Costs

Study of structural quality measures and maintenance effort across 20 customers in a large global system integrator



#### TQI increase of .24 decreased corrective maintenance effort by 50%

#### Reducing Operational Losses

Consortium for IT Software Quality



Large international investment bank Business critical applications



#### Percentage of variation in structural quality scores accounted for by various factors

Factor	# apps	Robust	Security	Perform	Change	Transfer
Technology	1606	20%	3%	16%	26%	7%

2017 CRASH Report Request from: info@castsoftware.com

## CISQ The Technical Debt Metaphor

**Technical Debt** — the future cost of fixing defects remaining in code at release, a component of the cost of ownership



Curtis, et al. (2012). IEEE Software.

#### **Tech. Debt by Quality Factor**

21%

45%

5% 4%

Cobol

29%

45%

5%

4%

Java-EE

24%

39%

9%

Mixed

Technologies

27%

42%

7%

4%

Oracle Server

24%

56%

Other



Curtis, et al. (2012). IEEE Software.

Consortium for IT Software Ouality


## Join CISQ ! www.it-cisq.org



## References



Anderson, D.J. (2010). Kanban. Sequim, WA: Blue Hole Press.

- Basili, V. Trendowicz, A., Kowalczyk, M., Heidrich, J., Seaman, C., Munch, J., & Rombach, D. (2014). *Aligning Organizations through Measurement*. Heidelberg: Springer.
- Boehm, B. (1987). Increasing software productivity. IEEE Computer, 20 (9), 43-57.

Boehm, B. et al. (2000). Software Cost Estimation with COCOMO II. Prentice-Hall.

Chrissis, M.B., Konrad, M., & Shrum, S. (2011). *CMMI Third Edition: Guidelines for Process Integration and Product Improvement*. Reading, MA: Addison Wesley.

Cohn, M. (2006). Agile Estimating and Planning. Upper Saddle River, NJ: Prentice-Hall.

Crosby, P. (1979). Quality Is Free. New York: McGraw-Hill.

Curtis, B., Sapiddi, J., & Szynkarski, A. (2012). Estimating the principle of an application's technical debt. *IEEE Software, 29* (6), 34-42.

Dion, R. (1993). Process improvement and the corporate balance sheet. IEEE Software, 10 (4), 28-35.

Duncker, R. (1992). *Presentation at the 25th Annual Conference of the Singapore Computer Society*. Singapore: November, 1992.

- Dyne, K. (1999). ESSI: Benefits from Continuous Software Improvements. In *Proceedings of ESEPG'99*. Milton Keyes, UK: ESPI Foundation.
- Flowe, R.M. & Thordahl, J.B. (1994). A Correlational Study of the SEI's Capability Maturity Model and Software Development Performance in DoD Contracts (AFIT/GSS/LAR/94D-2). Dayton, OH: Air Force Institute of Technology, Wright Patterson Air Force Base.



- Haley, T., Ireland, B., Wojtaszek, E., Nash, D., & Dion, R. (1995). Raytheon Electronic Systems Experience in Software Process Improvement: (Tech. Rep. CMU/SEI-95-TR-017).). Pittsburgh: Software Engineering Institute, Carnegie Mellon University.
- Ebert, C. & Dumke, R. (2007). Software Measurement. Berlin: Springer-Verlag.
- Fagan, E. & Davis, N. (2005). TSP / PSP at Intuit. *Proceedings of the TSP 2005*. Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University.
- Fenton, N. (2006). New directions for Software Metrics. CIO Symposium on Software Best Practices.
- Garmus, D. & Herron, D. (2001). Function Point Analysis. Boston: Addison-Wesley.
- George, E. (2005). Project management with TSP. *Proceedings of the TSP 2005*. Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University.
- Herbsleb, J, Carleton, A., Rozum, J., Siegel, J., & Zubrow, D. (1994). Benefits of CMM-Based Software Process Improvement: Initial Results (Tech. Rep. CMU/SEI-94-TR-13). Pittsburgh: Software Engineering Institute, Carnegie Mellon University.
- Humphrey, W. (1997). Introduction to the Personal Software Process. Boston: Addison Wesley.
- Humphrey, W. (2000). Introduction to the Team Software Process. Boston: Addison Wesley.
- Humphrey, W. (2006). TSP: Leading a Development Team. Boston: Addison Wesley.
- Jackson, D. (2009). A direct path to dependable software. Communications of the ACM, 52 (4).
- Kan, S.H. (1995). Metrics and Models in Software Quality Engineering. Reading, MA: Addison-Wesley.
- Keeni, B. (2000). IEEE Software, 16 (4).



Larman, C. (2004). Agile and Iterative Development. Boston: Addison Wesley.

Lydon, T. (1995). Productivity drivers: Process and capital. In *Proceedings of the 1995 SEPG Conference*. Pittsburgh: Software Engineering Institute, Carnegie Mellon University.

McGarry, J. et al. (2002). Practical Software Measurement. Reading, MA: Addison Wesley.

Paulk, M., Weber, C., Curtis, B., & Chrissis, M. (1995). *The Capability Maturity Model: Guidelines for Improving the Software Process*. Reading, MA: Addison-Wesley.

Peterman, W. (2000). IEEE Software, July/August issue.

- Poppendieck, M. & Poppendieck, T. (2007). *Implementing Lean Software Development*. Boston: Addison-Wesley.
- Sakamoto, K., Kishida, K., & Nakakoji, K. (1996). Cultural adaptation of the CMM. In Fuggetta, A. & Wolf, A. (Eds.), Software Process. Chichester, UK: Wiley, 137-154.
- Sinha, M. (2006). Performance models Enabling organizations prediction capability. *Proceedings of SEPG'06*. Pittsburgh: Software Engineering Institute, Carnegie Mellon University.

Schwaber, K. & Beedle, M. (2002). Agile Software Development Using Scrum.

Spinellis, D. (2006). Code Quality. Addison-Wesley

- Van Epps, S. & Marshall, R. (2012). Using TSP at the end of the development cycle. *Proceedings of TSP 2012*. Pittsburgh: Software Engineering Institute, Carnegie Mellon University.
- Vu. J.D. (1996). Software process improvement: A business case. In *Proceedings of the European SEPG Conference*. Milton Keynes, UK: European Software Process Improvement Foundation