NGOMS Metrics Workshop
Aug 31, 2021

Peter Capell, Nanette Brown, Keith Korzec

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

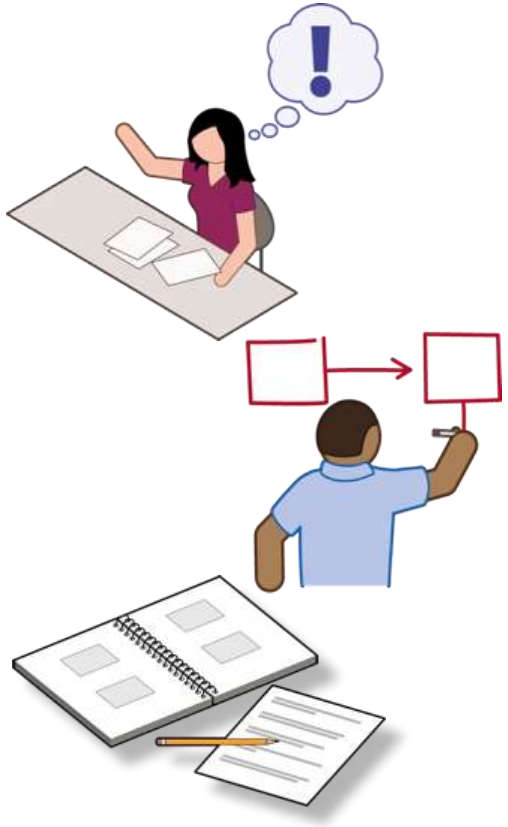
DM21-0754

Agenda for Today

- Welcome & Introductions
- Setting the Context
  - Build User Stories (Roles and Purpose)
  - Next Steps
  - Briefing Metrics
  - Retrospective

Welcome and Introductions

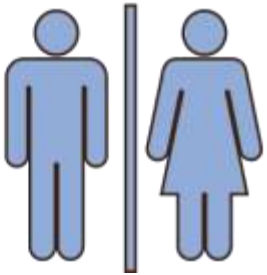
Approach



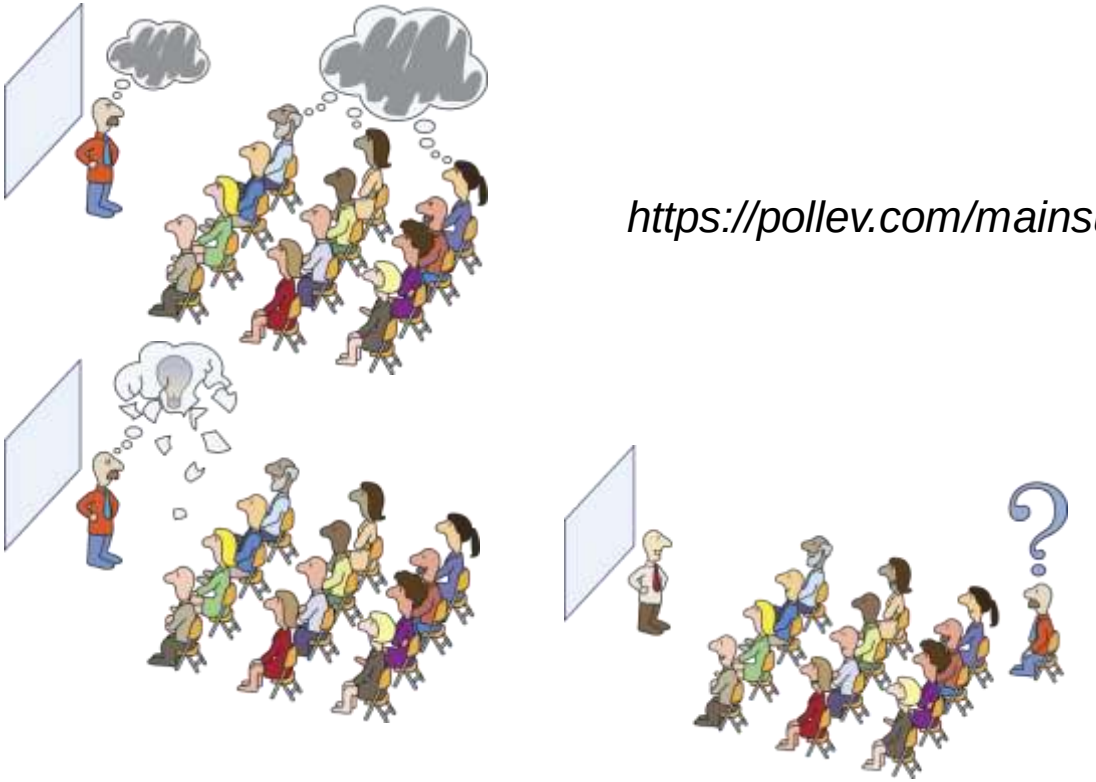
Mix of

- DISCUSSIONS and presentations
- Exercises

Course Logistics



Who we are — Who are You?



https://pollev.com/mainsummit799

Setting the Context

Setting the Context

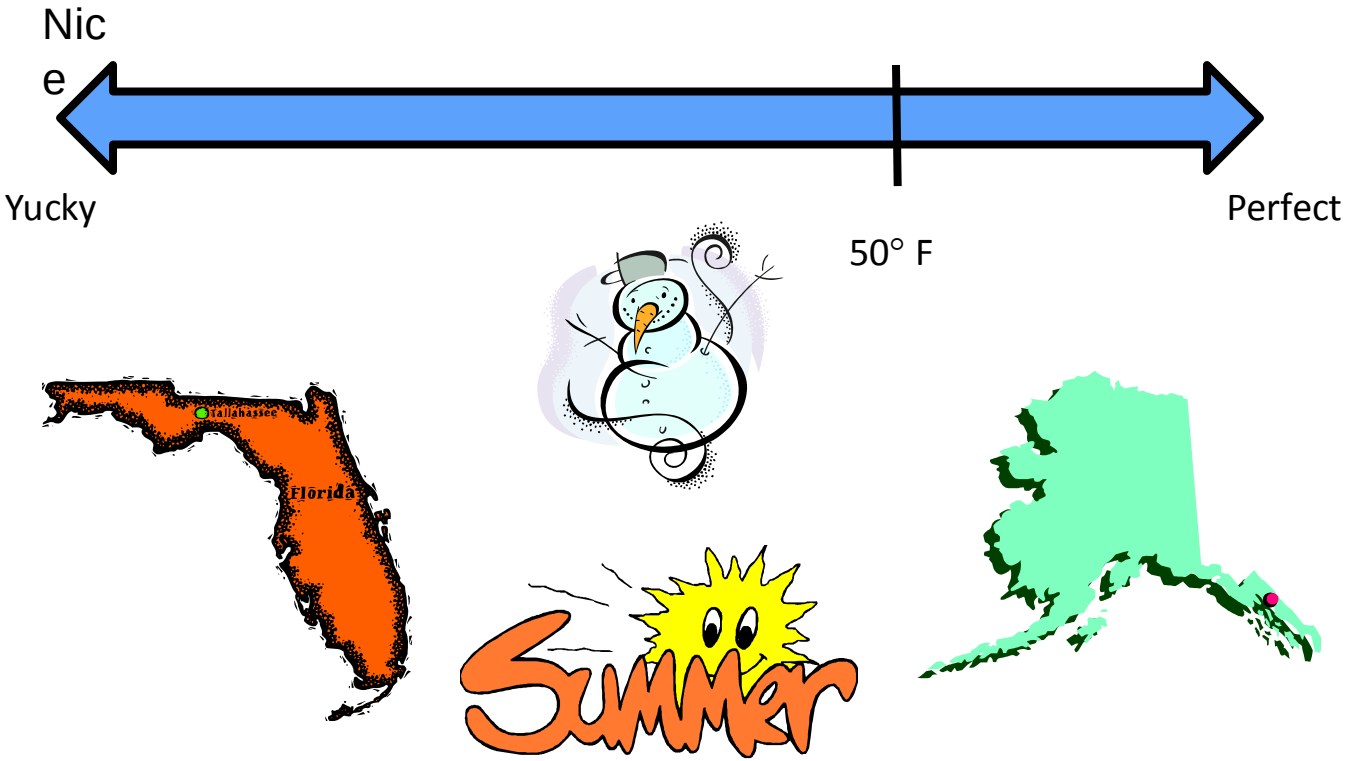
What do you think of when you hear the work “METRICS”?

Everybody is talking about metrics & measures – I'm confused. Aren't they the same thing?

Well .....

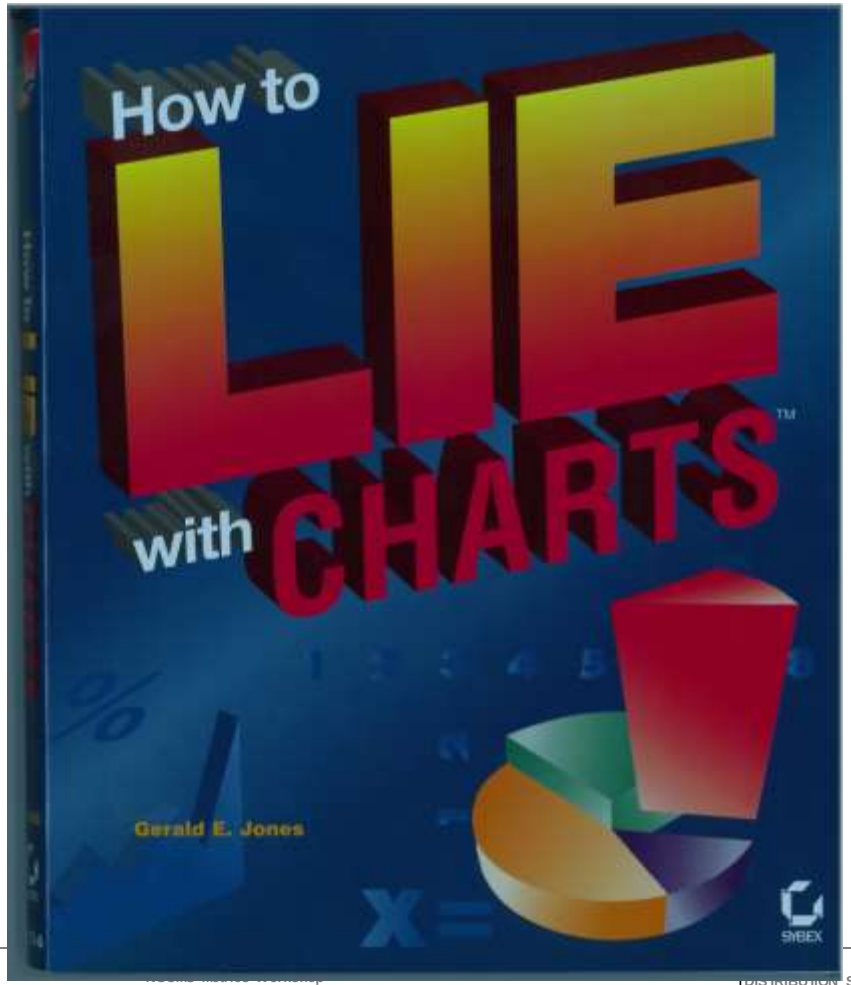
Managing By Metrics

How's the temperature outside?



A measurement or metric program will NOT guarantee a successful program

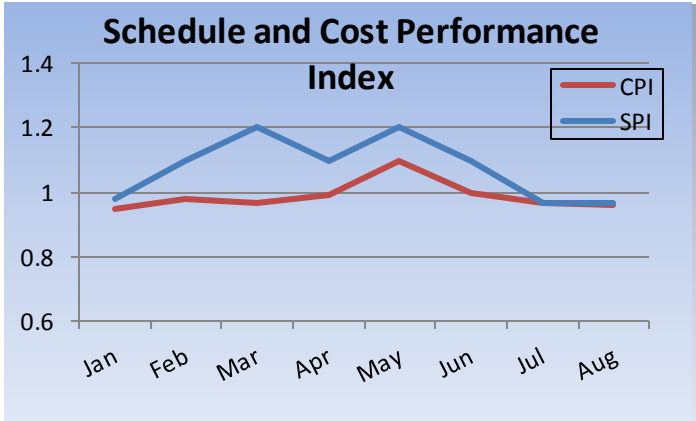
58.3% of statistics are made up

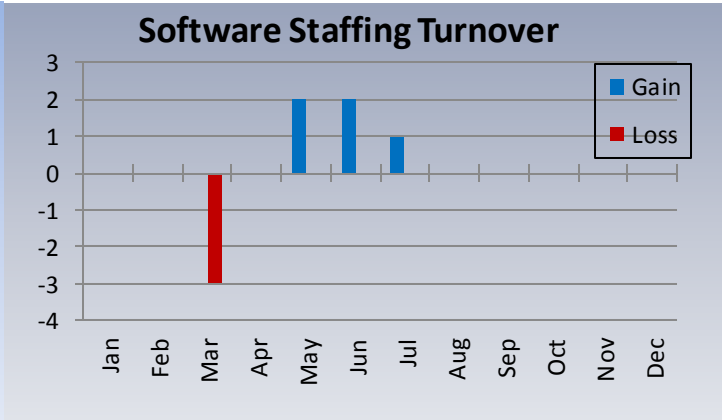


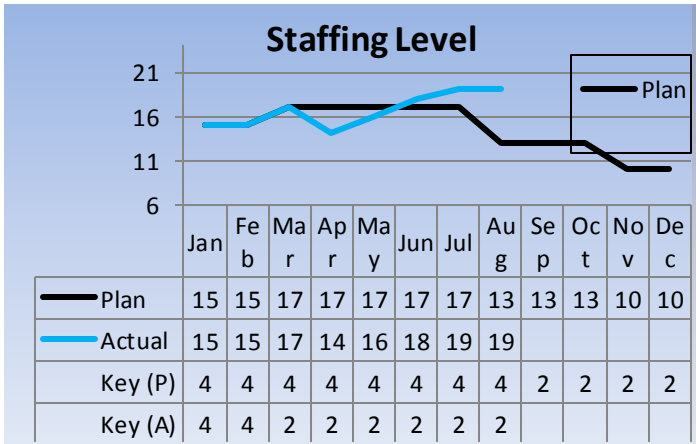
Measurement by itself does not control or improve. Use measurement to gain insight to objectively plan, manage and communicate

Metrics in Practice

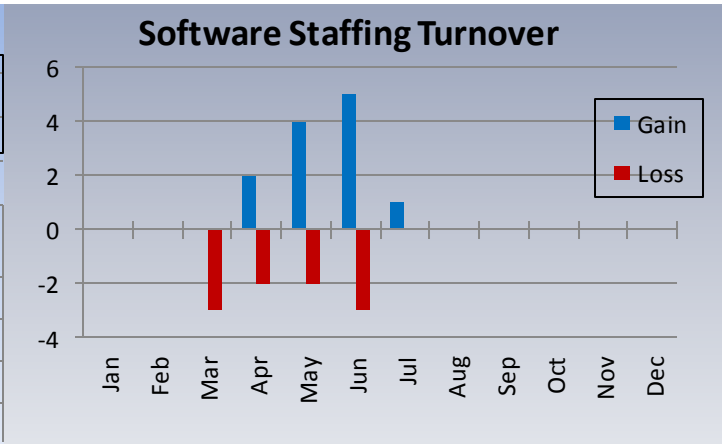
What do all those squiggly lines mean?







	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Plan	15	15	17	17	17	17	17	13	13	13	10	10
Actual	15	15	17	14	16	18	19	19				
Key (P)	4	4	4	4	4	4	4	4	2	2	2	2
Key (A)	4	4	2	2	2	2	2	2				



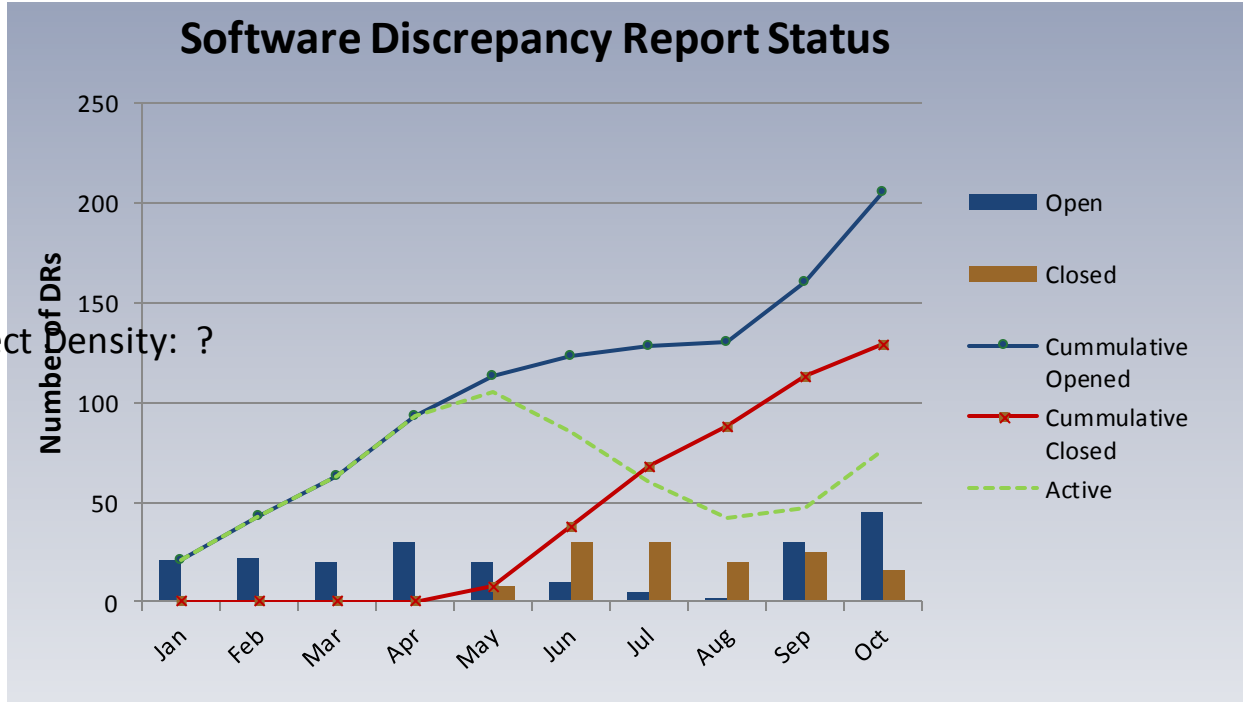
Rarely does one measure give you the whole picture.

Using multiple indicators ensure you are accurately interpreting your data.

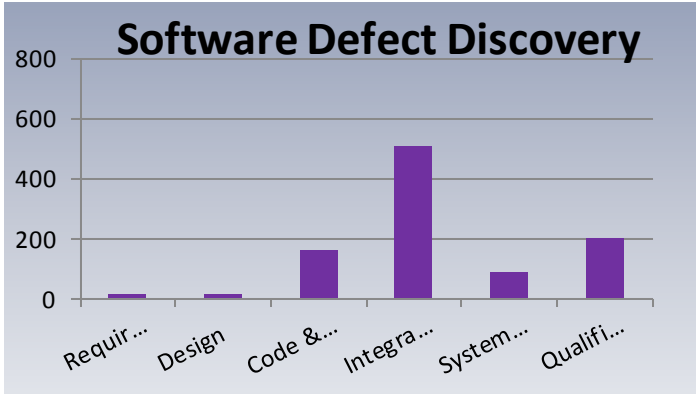
It DEPENDS!

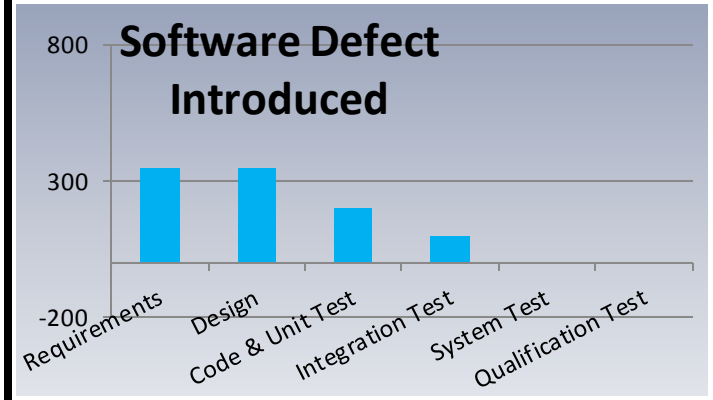
Defect Density: ?

Discrepancy Report Close Ratio: 63%



Defect Density: ?



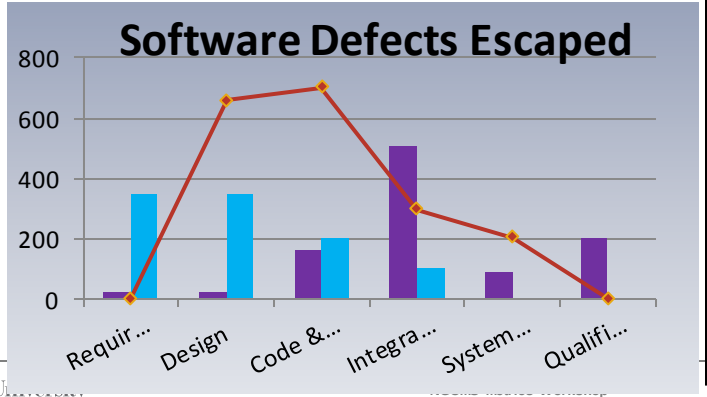


Discovery

Injection

Escape

Cost



	Introduced	Found	Cost for removal
Requirements & Design	70%	3.5%	1X
Code & Unit Test	20%	16%	5X
Integration Test	10%	50.5%	16X
System Test	0	9%	40X
Acceptance Test	0	20.5%	110X

System Type	Average Defect Density
Command and control	**0.106**
Command, control and communications	**0.011**
Military ground vehicle	0.106
Satellite	0.087
Large stationery capitol equipment	0.649
GPS	0.134
Power systems	1.093

[SoftRel, LLC 2007]

Commercial SW	Size (millions)	Measured Defect Density
Linux 2.6	7	0.62
Red Hat Linux 6.2	17	0.12329
Red Hat Linux 7.1	30	0.12597
Windows 95	15	0.3333
Windows 98	18	0.5556
Windows XP	40	2.6625
Windows NT 4.0	16	0.625
Win 2000	35	1.80

[Alhazmi OH et al., Computers & Security 2006]

Even with such a low DD, a system like NOMS, with 1 million lines, can expect to field between 11 and 106 defects

Build User Stories

Defining Needs

Describe YOUR SAFe / Agile role.

Use yellow stickies

What decisions do you make?

What do you NEED TO KNOW to do your job?

Use orange stickies

Example Schedule and Progress Decisions

Typical Questions:

- Is the project meeting scheduled milestones?
- Are delivery dates slipping?
- How are specific activities and products progressing?
- Is capability being delivered as scheduled in incremental builds and releases?

Defining Needs

Describe YOUR SAFe / Agile role.

Use yellow stickies

What decisions do you make?

Use orange stickies

What data do you use now? What information would make it easier?

Use pink stickies

Example Schedule and Progress Data Sources

- ☑ Milestone Performance
  - Milestone dates
- ☑ Work Unit Progress
  - Component Status
  - Requirement Status
  - Test Case Status
  - Problem Report Status
  - Review Complete
  - Change Request Status
- ☑ Incremental Capability
  - Build Content - Component
  - Build Content - Function

Adapted from[PSM2003]

Board Layout

Re-arrange Board

Identify Source(s) of Data & Finalize Metrics

Where can we get the data you need?

How often will you use it?

Automated? Dashboard? Report?

Use green stickies

Next Steps

Identify Owner

- Decompose data source (if necessary)
- Who will track down the data source for each of these?
- Create tickets in Jira

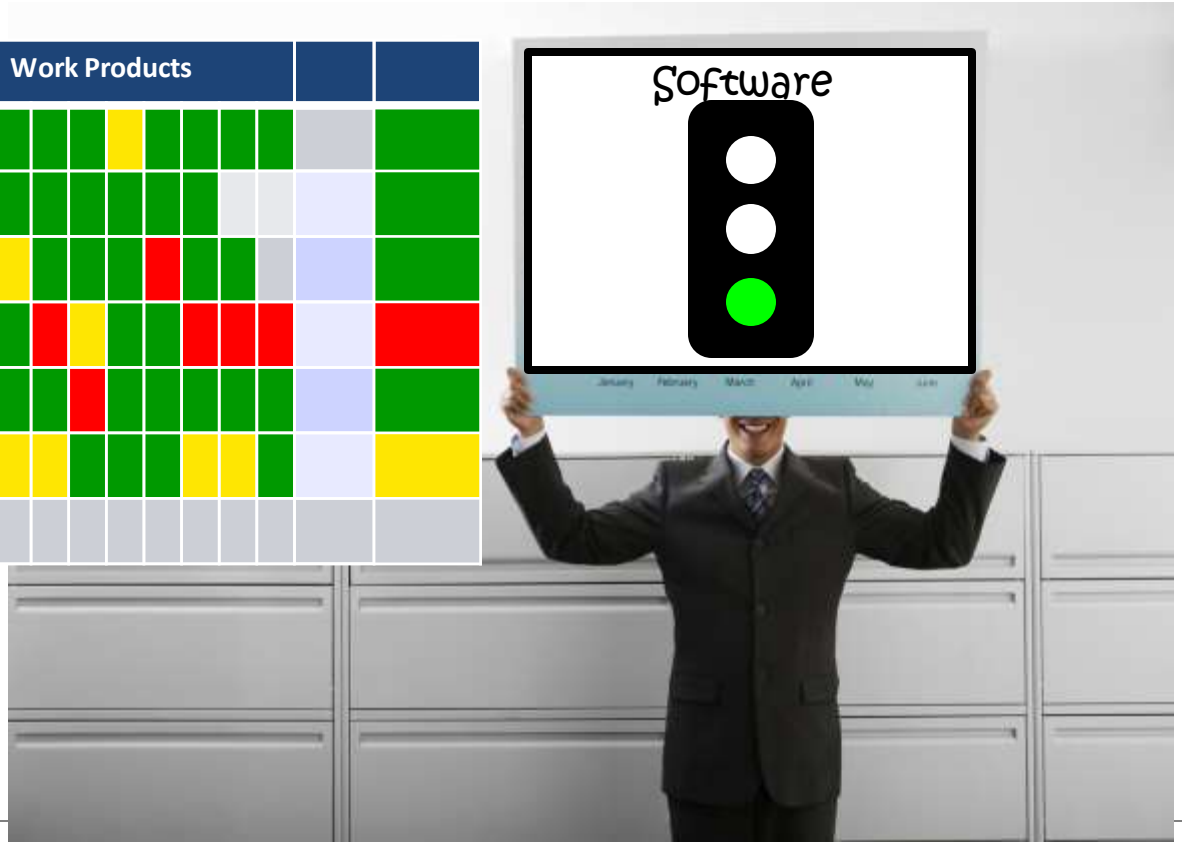
Briefing Metrics

When briefing: make the complex understandable

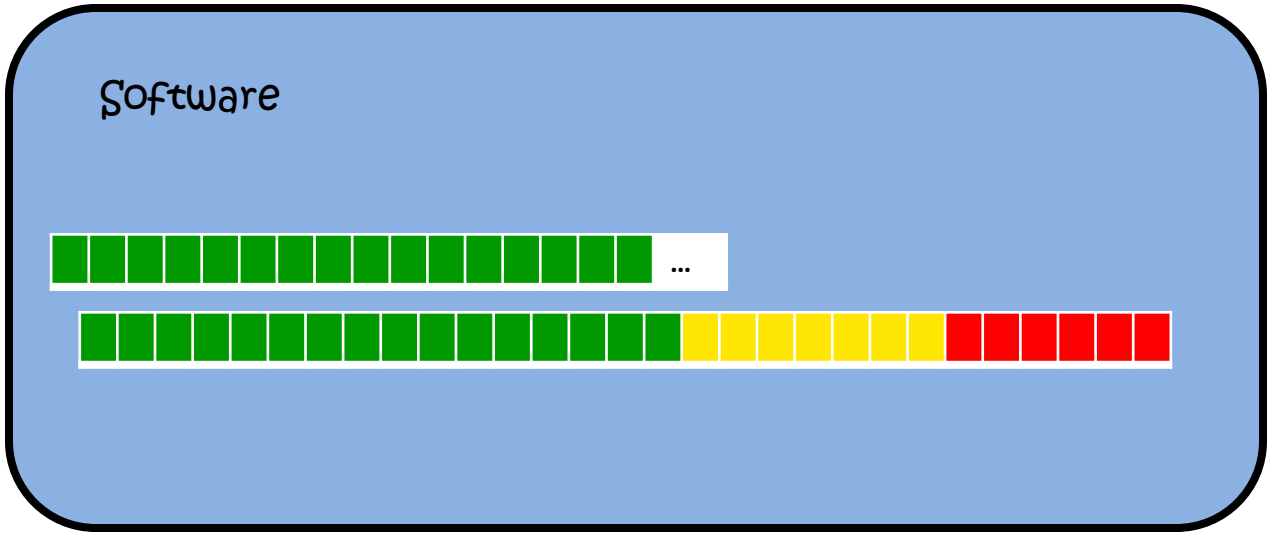
“Everything should be made as simple as possible, but not simpler” - Albert Einstein

Measurement Roll-Up

	Work Products		
Team 1			
Team 2			
Team 3			
Team 4			
Team 5			
Team 6			



Measurement Roll-Up



“Roll-ups” have the potential of hiding real issues and risks.

Summary

Crawl, Walk, Jog, Run

You can’t get to perfection overnight

Things to Remember

How do you know what measures to collect?

- Support your business goals
- Provide insights into the management issues that are most important to you
- Data-collection activities are better able to stay focused on their intended objectives

How do you know when you need the measures?

- Experience and historical data!

Things to Remember (cont.)

What do you look at?

- It is OK to examine detailed data early in the project
- Remember the “big picture”
- Partial data is better than no data
- This could have led to an earlier troubleshooting effort and earlier detection of root causes

How much variance is “enough” to warrant action?

- As experience with data increases, past projects will serve as benchmarks

You may get misleading or incomplete data.

The Program Office must be knowledgeable enough to assess the measurement data provided by the contractor

Retrospective

Backup Slides

Identifying the needs for metrics

Use a workshop to identify job-critical measures and document in a familiar agile format.

Examples:

- As a member of the ELT I want metrics on capability progress so that I can accurately report progress.
- As the program manager, working on XXX release, I want to know the likelihood of a given feature slipping into the next PI, so that I can assess the risk to the timely completion of work to mature and deliver the capability which relies on this feature.
- As a systems engineer I need measures of technical debt and whether our unresolved DRs are increasing or decreasing after each PI so that I can adjust my plans on new work for future PIs
- As a test engineer, I need to understand the content of the next build, so that I can assure the availability of needed test assets and timely integration testing.

Example: Schedule and Progress

The completion of program milestones, significant events, and individual work items

Typical Questions:

- Is the project meeting scheduled milestones?
- Are delivery dates slipping?
- How are specific activities and products progressing?
- Is capability being delivered as scheduled in incremental builds and releases?

Typical Measurements

- ☑ Milestone Performance
  - Milestone dates
- ☑ Work Unit Progress
  - Component Status
  - Requirement Status
  - Test Case Status
  - Problem Report Status
  - Review Complete
  - Change Request Status
- ☑ Incremental Capability
  - Build Content - Component
  - Build Content - Function

Adapted from[PSM2003]

Example: Resources and Cost

The relationship between work to be performed and resources assigned and used

Typical Questions:

- Is effort being expended according to plan?
- Is there enough staff?
- Is the proper mix of staff assigned?
- Is the project spending meeting budget and schedule objectives?
- Are necessary facilities and equipment available as planned?

Typical Measures

- ☑ Personnel
  - Effort
  - Staff Experience
  - Staff Turnover
- ☑ Financial Performance
  - Earned Value
  - Cost
- ☑ Development Environment
  - Resource Availability
  - Resource Utilization

Adapted from [PSM2003]

Example: Product Size and Stability

The quantity and volatility of delivered software or functionality

Typical Questions:

- How big is the software?
- How much change is occurring?
- Is there an excessive amount of requirements change?

Typical Measures

☑ Physical Size and Stability
- System Element Trends
- Interface Complexity
- Lines of Code Trends

☑ Functional Size and Stability
- Requirements Trends
- Architecture Element Trends
- Work Unit Backlog Size Trends
- Function Points

Adapted from [PSM2003]

Example: Product Quality

The ability of the software to meet the users' needs

Typical Questions:

- Is the product good enough for delivery to the user?
- Are identified problems being resolved?
- Can functionality be re-hosted on different platforms?
- How many vulnerabilities are identified and remediated by phase?
- Is the user interface adequate and appropriate for operations?
- Are failure rates within acceptable bounds?

Typical Measures

- ☑ Functional Correctness
  - Defect Profiles
  - Defect Density
- ☑ Supportability - Maintainability
  - Time to Restore
  - Cyclomatic Complexity
- ☑ Portability
  - Interface Compliance
- ☑ Security
  - Profile of Vulnerabilities

Adapted from [PSM2003]

Example: Process Performance

The relationship between the capability of the contractor and the needs of the program

Typical Questions:

- How consistently does the project implement its defined processes?
- Are the processes efficient enough to meet current commitments?
- How much re-work is occurring?

Typical Measures

☑ Process Compliance
- Process Audit Finding Distribution
- Process Reference Maturity/Capability Rating

☑ Process Efficiency
- Productivity Performance Trends
- Cycle Time Performance Trends

☑ Process Effectiveness
- Defect Containment
- Test Coverage
- Rework Effort Distribution

Adapted from [PSM2003]

Example: Technology Effectiveness

The maturity and viability of the proposed technical approach

Typical Questions:

- Can technology meet all requirements?
- Is the technology ready to be used?
- Does new technology pose a risk due to too many changes?
- Will COTS software mitigate some risks?

Typical Measures

- ☑ Technology Suitability
  - Requirements Coverage
- ☑ Technology Maturity
  - Technology Maturity Trends
- ☑ Technology Volatility
  - Technology Baseline Changes Trends

Adapted from [PSM2003]

Example: Customer Satisfaction

The relationship between the system delivered to the expectations of the customer

Typical Questions:

- Is the program meeting user expectations?
- How quickly are customer issues being addressed?

Typical Measures

- ☑ Customer Feedback
  - Satisfaction Ratings Trends
  - Award Fee Distributions

- ☑ Customer Support
  - Support Request Distributions
  - Support Time Trends

Adapted from [PSM2003]