



National Defense Industrial Association

A Path Toward Consensus Measures for Iterative Software Development

An industry/DoD collaboration

Cheryl Jones (US Army)

Geoff Draper (L3Harris)

NDIA Systems and Mission Engineering Conference, Oct 2019

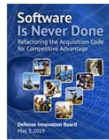
Overview – SW Measurement Framework



DSB SW

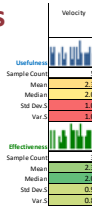


DIB SWAP

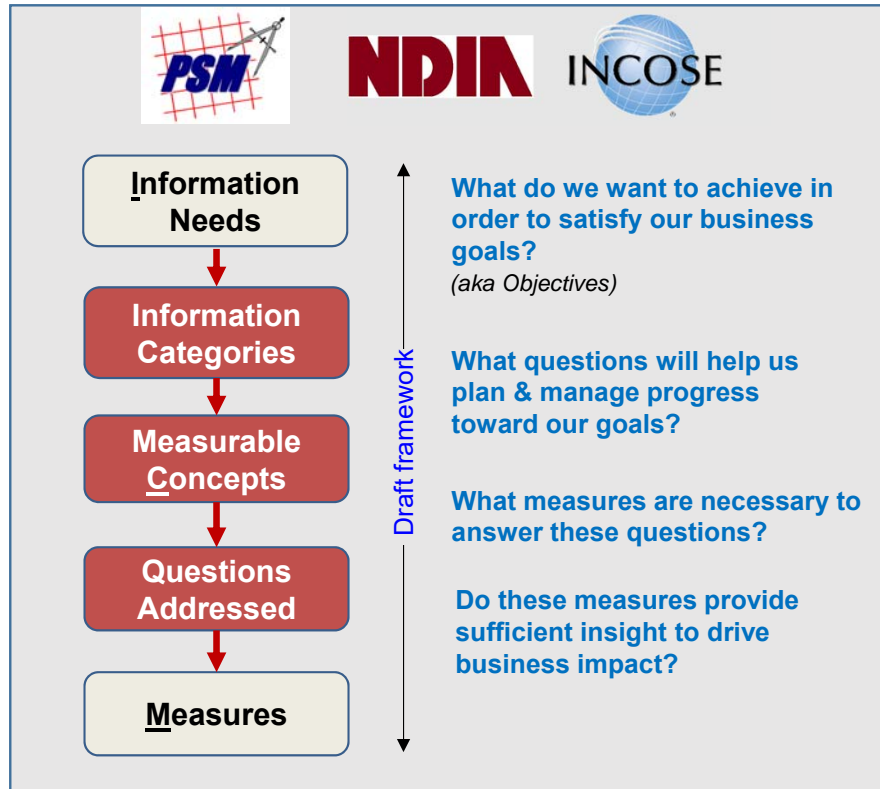


Info Needs
Measures
Surveys

- PSM
- NDIA
- INCOSE
- SERC



Draft DoD
SW policy

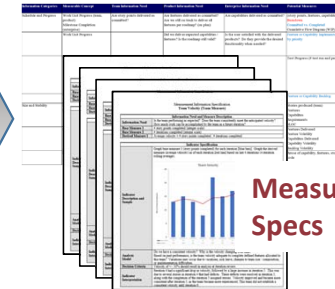


PSM, NDIA, and INCOSE are collaborating on development of a consensus industry measurement framework for agile/CID

<https://www.ndia.org/divisions/systems-engineering/studies-and-publications>

NDIA WG recommendations: DSB #3 (measures)

Picture of Success (end state)	
Consensus frameworks	<ul style="list-style-type: none"> Objectives first - measures aligned and tailored from information needs, goals and constraints, at program and enterprise levels
Modernized measures	<ul style="list-style-type: none"> Migration toward consensus alternatives to traditional waterfall and phase-based SW measures (LOC, EVM, milestones, ...) Derived from SW factory processes, automated by toolchain Basis for measuring cost and schedule vs. plan
History-based estimates	<ul style="list-style-type: none"> Repositories collect performance-based measures (e.g., WBS, staff, cost, productivity) supporting future comparisons, basis of estimates, proposals, and program monitoring



ICM Table

- Information Categories
- Measurable Concepts
- Information Need (team, product, enterprise)
- Potential measures

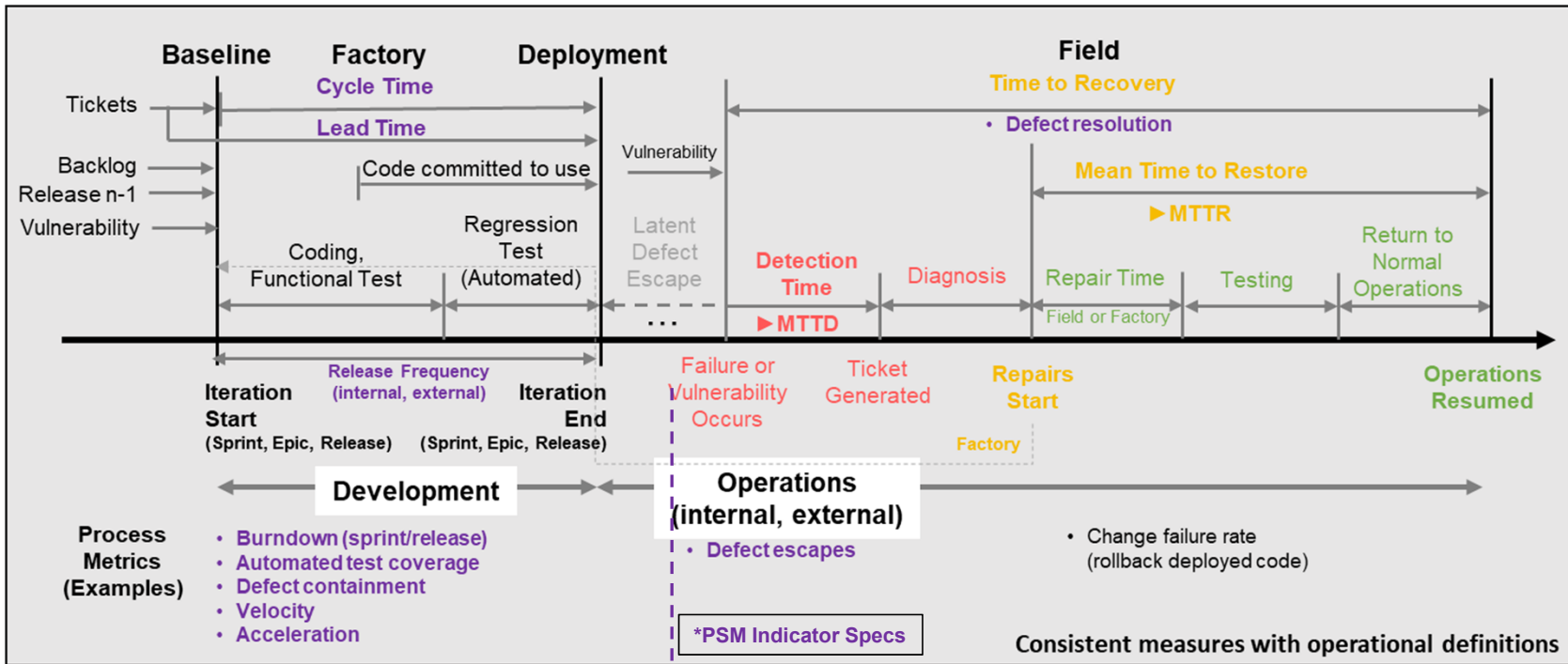
Measurement Specs

- Information Need
- Measures (base, derived)
- Indicator description, sample
- Analysis model
- Decision criteria
- Interpretation, guidance
- Implementation considerations

Acceleration	Automated Test Coverage	Burndown (Sprint/Release)	Defect Containment
Defect Escapes	Defect Resolution	Cycle Time	Lead Time
Release Frequency	Velocity	...	

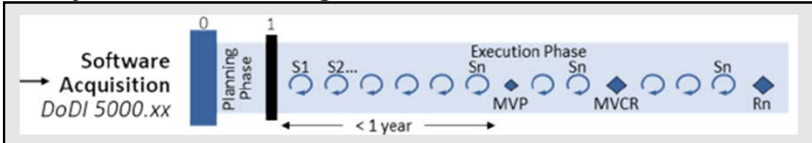
NDIA/INCOSE/PSM Continuous Iterative Development and Sustainment WG

Aligning the PSM framework and measures with DoD SW policy and enterprise improvement



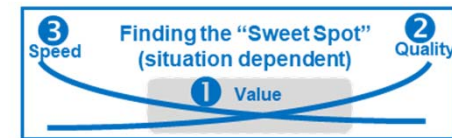
We will be asking for your input on which measures you find most useful and effective in your organizations

Policy, Guidance, Oversight



Measures, goals, and priorities are tailored and aligned based on objectives and information needs

Program ► Product ► Enterprise



References:

- Defense Science Board, Design and Acquisition of Software for Defense Systems, Feb 2018
- Defense Innovation Board Metrics for Software Development, version 0.9, 9 Jul 2018
- MTTR, MTBF, or MTTF? A Simple Guide to Failure Metrics. <https://limblecmms.com/blog/mttr-mtbf-mttf-guide-to-failure-metrics/>

PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

ICM Table (Draft) Excerpts most relevant to initial PSM agile measurement framework – 1 of 3

Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need	Potential Measures
Schedule and Progress	Work Unit Progress (team, product) Milestone Completion (enterprise)	Are story points delivered as committed?	Are features delivered as committed? Are we still on track to deliver all features per roadmap? (on plan)	Are capabilities delivered as committed?	(story points, features, capabilities) Bumdown Committed vs. Completed Cumulative Flow Diagram (WIP)
	Work Unit Progress		Did we deliver expected capabilities / features? Is the roadmap still valid?	Is the user satisfied with the delivered products? Do they provide the desired functionality when needed?	Feature or Capability Implementation by priority
	Work Unit Progress		Is the integration and test progress proceeding as planned?		Test Progress (# test run and passed)
	Work Backlog		How much outstanding technical or mission debt exists?		Feature or Capability Backlog
Size and Stability	Functional Size and Stability Physical Size and Stability	How big is our system?	How big is our system?	How big is our system?	Stories produced (team) Features Capabilites Requirements SLOC
	Functional Size and Stability		How volatile are capabilities or features? Are we adding more features? What is the ability to accommodate changes in customer desirements?	How volatile are capabilities or requirements? What is the ability to accommodate changes in customer desirements?	Features Delivered Feature Volatility Capabilites Delivered Capability Volatility Backlog Volatility
	Functional Size and Stability	How much of the product is newly developed vs. reused from other sources?			Reuse of capability, features, stories, code

ICM Table (Draft) Excerpts most relevant to initial PSM agile measurement framework – 2 of 3

Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need	Potential Measures
Product Quality	Functional Correctness	Does new code functionality work as expected?	Does new code functionality work as expected?	Is rework identified and managed?	Stories Accepted (increment demo) Rework Stories Change Reports (defects) Written
	Functional Correctness	Does new code break previous functionality?	Does new code break previous functionality? (change failure rate, rollback)		Change reports (defects) written Rework hours Change Failure Rate or Defect Density
	Functional Correctness		How many defects escape the increment?		Defects Found in Pipeline (saves)
	Functional Correctness		What is the quality of code deployed to the field?	What is the quality of code deployed to the field?	Defect Escapes to field Defect Escape Ratio
	Security - Safety		How secure is the product		Vulnerabilities by severity
	Supportability - Maintainability Dependability - Reliability		What is the reliability and availability of operational service capabilities?		Mean-Time-To: MTTD (Detect) MTTR (Repair or Rest ore) MTBF (Between Failure) MTTF (Failure) Ao (Operational Availability)

PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

ICM Table (Draft) Excerpts most relevant to initial PSM agile measurement framework – 3 of 3

Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need	Potential Measures
Process Performance (Process Effectiveness)	Process Efficiency - Speed Security - Safety		How quickly can new security vulnerabilities be patched and deployed to fielded products?		Security vulnerability lead time Mean Time to Restore
	Process Efficiency - Speed Supportability - Maintainability Dependability - Reliability		How quickly can we address bug reports from the field?		Mean Time to Restore MTTD
	Process Efficiency - Speed	Is the team performing as expected?	Are teams performing as expected?		Velocity (average story points per increment) Capacity (staffhours per increment) Story points delivered vs. committed (on average) Cumulative flow diagrams
	Process Efficiency - Speed		How long does it take to deploy an identified feature/capability?		Lead time
	Process Efficiency - Speed		What is the frequency of product release or deployment?	What is the frequency of product release or deployment?	Release or deployment frequency
	Process Efficiency - Speed	How long does it take to release a viable product?	How long does it take to release a viable product?	How long does it take to release a viable product?	Release frequency Cycle time (increment, release, mean/median) Time to Minimum Viable Product (MVP)
	Process Efficiency - Speed		How much time does it take to conduct a full regression test? How much time for the automated regression test?		Test duration Automated test duration
	Process Effectiveness		How much of the testing is automated? How often do we perform automated testing?	How much of the testing is automated? How often do we perform automated testing?	Automated test coverage Automated test frequency
	Process Effectiveness	Is the backlog being managed appropriately?	Is the backlog being managed appropriately?		Cumulative flow diagram Backlog readiness
Customer Satisfaction	Customer Support		How long does it take to get a viable product released? (specific)	How long does it take to get a viable product released? (multiple systems) - time to market	Time to Minimum Viable Product (MVP)

PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

We would like your feedback on the draft framework for continuous iterative development

- **Information needs** – deferred due to time constraints; see backup charts
- **Candidate measures**

Two separate evaluations are desired:

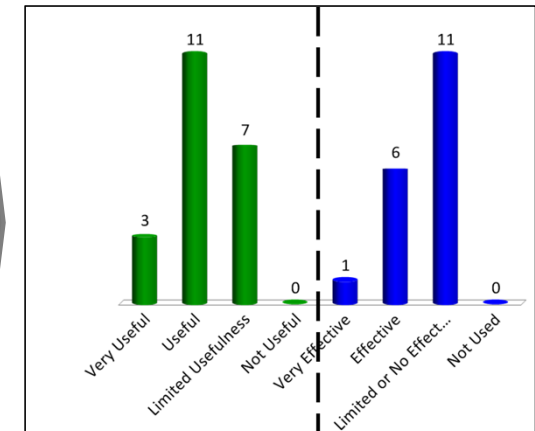
- **Usefulness:** Is the measure itself useful for providing insight?
- **Effectiveness:** How effectively does your organization use it?

Example:

Is [measure name] a **useful** measure, and how **effectively** is it used to provide insight and impactful action in your organization?

- | | | | |
|-------------------|----|-----------------------------|-------------------------------|
| ↑
Insight
↓ | 1. | Very Useful | Select 1
from here |
| | 2. | Useful | |
| | 3. | Limited Usefulness | |
| | 4. | Not Useful | |
| ----- | | | |
| ↑
Impact
↓ | 5. | Very Effective | and 1
from here |
| | 6. | Effective | |
| | 7. | Limited or No Effectiveness | |
| | 8. | Not Used | |

THEORY PRACTICE



Your feedback will help us validate and improve the draft PSM framework for government/industry consensus and potential use in defense software acquisition

How would you best characterize your organization?

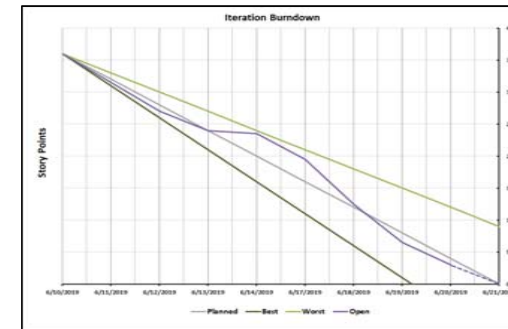
- 1. U.S. Government (DoD, agency)**
- 2. U.S. Defense Industry**
- 3. Academia / FFRDC**
- 4. Commercial Industry**
- 5. Other**



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

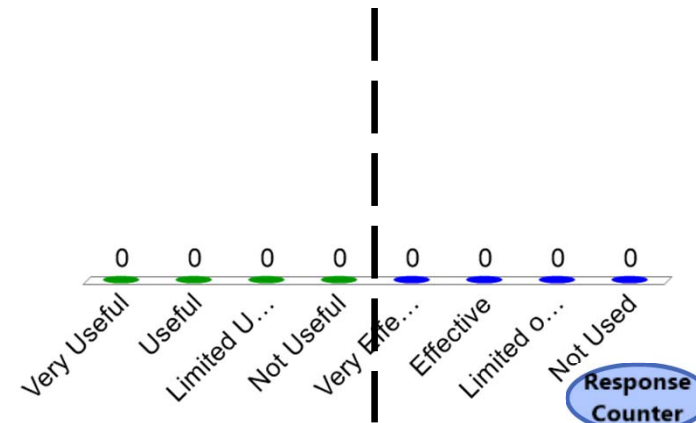
Is **Sprint Burndown** a useful measure, and how effectively is it used to provide insight and impactful action in your organization?

Information Need and Measure Description	
Information Need	What is the status of the iteration? Will all the remaining committed work be completed by the end of the iteration? Will the team deliver the committed story points?
Base Measure 1	Planned story points (integer scale)
Base Measure 2	Completed story points (integer scale)
Derived Measure 1	Open story points = planned story points – completed story points



- ↑ Insight
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 - ↓ 4. **Not Useful**
- ↑ Impact
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 - ↓ 8. **Not Used**

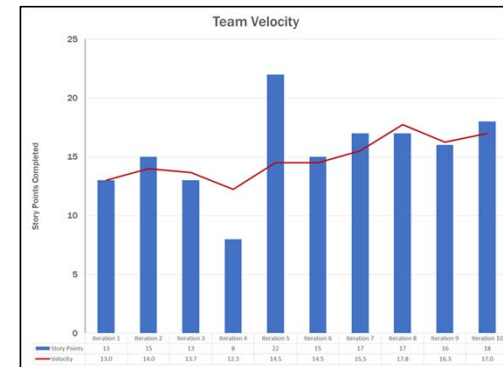
(select 1 from each)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

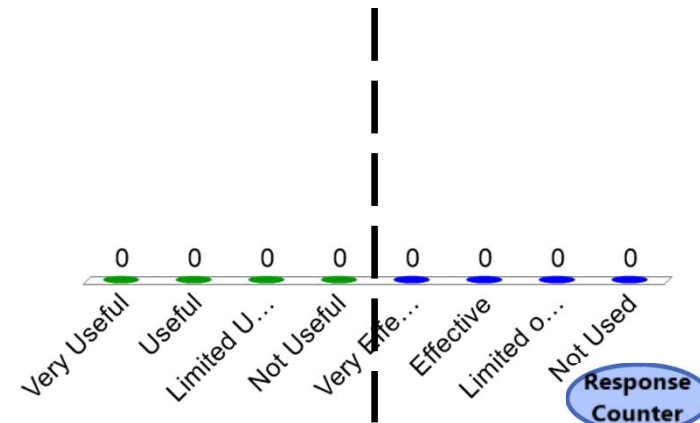
Is **Velocity** a **useful** measure, and how **effectively** is it **used** to provide insight and **impactful** action in **your** organization?

Information Need and Measure Description	
Information Need	Is the team performing as expected? Does the team consistently meet the anticipated velocity? How much work can be accomplished by the team in a future iteration?
Base Measure 1	# story points completed (integer scale)
Base Measure 2	# iterations completed (integer scale)
Derived Measure 1	Average velocity = # story points completed / # iterations completed



- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 4. **Not Useful**
- ↓ **Impact**
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 8. **Not Used**

(select 1 from each)

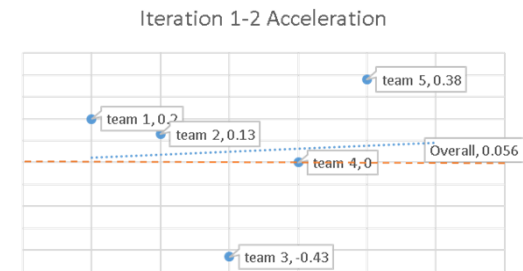


PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

Is **Acceleration** a **useful** measure, and how **effectively** is it used to provide insight and impactful action in **your** organization?

	Iter 1 pts	Iter 2 pts	Acceleration
team 1	10	12	0.2
team 2	8	9	0.13
team 3	14	8	-0.43
team 4	12	12	0
team 5	8	11	0.38
Overall			0.056

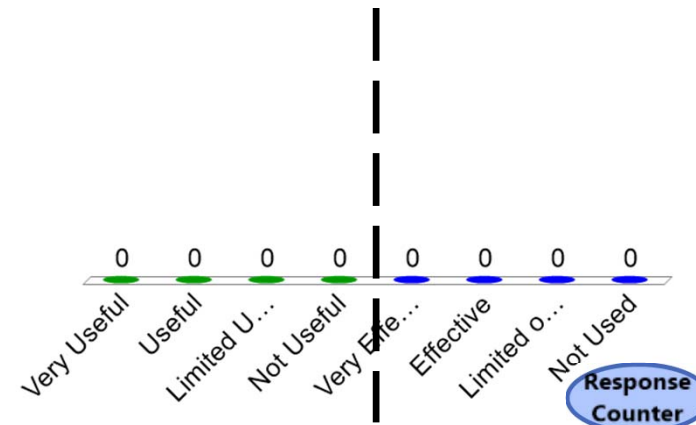
Sample calculation:
Team 1 acceleration = $12 - 10 / 10 = .2$
(20% positive acceleration)



Information Need and Measure Description	
Information Need	Is the team's productivity increasing, decreasing or holding steady?
Base Measure 1	# story points completed (velocity) this increment (integer scale)
Base Measure 2	# story points completed (velocity) in previous comparison increment (integer scale)
Derived Measure 1	Team Acceleration = $(\text{Current increment velocity} - \text{comparison increment velocity}) / \text{comparison increment velocity}$
Derived Measure 2	Overall Acceleration = $\text{Team Acceleration 1} \dots \text{Team Acceleration N} / \text{N}$

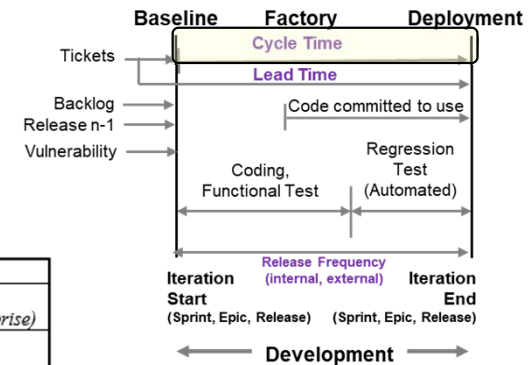
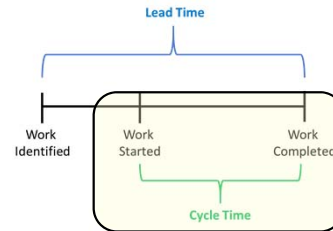
- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 4. **Not Useful**
- ↓ **Impact**
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 8. **Not Used**

(select 1 from each)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

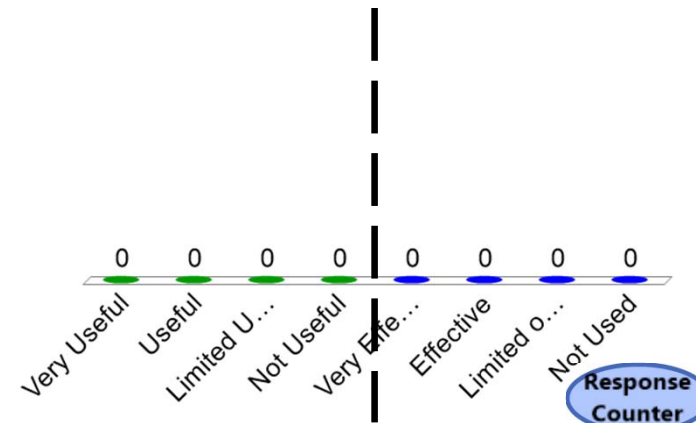
Is **Cycle Time** a **useful** measure, and how **effectively** is it used to provide insight and impactful action in **your** organization?



Information Need and Measure Description	
Information Need (Cycle Time)	How long does it take to complete a process activity? (team) How long does it take to develop an identified feature/capability or release? (product or enterprise)
Information Need (Lead Time)	How long does it take to get a viable product released (time to market)?
Base Measure 1	Start time for a process activity (date and time)
Base Measure 2	End time for a process activity (date and time)
Derived Measure 1	Elapsed Time = $\text{ceiling}(\text{End Time} - \text{Start Time})$ (Units may vary based on team context, capability, cadence; hours, days, weeks, months. May also vary based on calendar time vs. work days. Results with fractional values are rounded up to the next unit.)

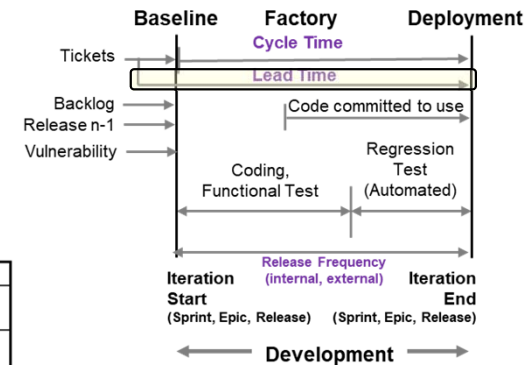
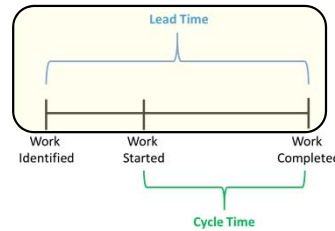
- ↑ Insight
1. Very Useful
 2. Useful
 3. Limited Usefulness
 - ↓
 4. Not Useful
- ↑ Impact
5. Very Effective
 6. Effective
 7. Limited or No Effectiveness
 8. Not Used

(select 1 from each)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

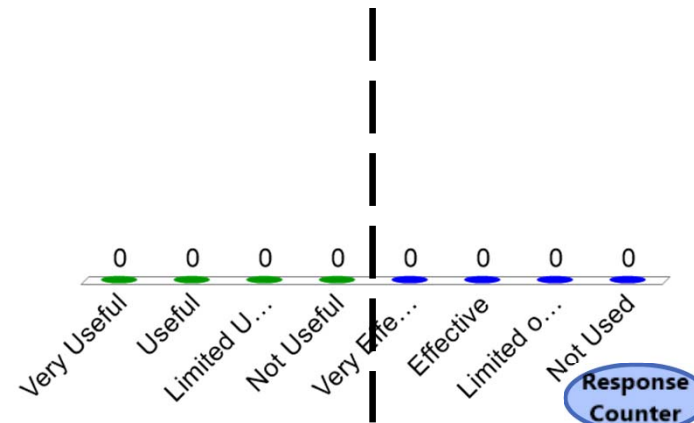
Is **Lead Time** a **useful** measure, and how **effectively** is it used to provide insight and impactful action in **your** organization?



Information Need and Measure Description	
Information Need (Cycle Time)	How long does it take to complete a process activity? (team) How long does it take to develop an identified feature/capability or release? (product or enterprise)
Information Need (Lead Time)	How long does it take to get a viable product released (time to market)?
Base Measure 1	Start time for a process activity (date and time)
Base Measure 2	End time for a process activity (date and time)
Derived Measure 1	Elapsed Time = $\text{ceiling}(\text{End Time} - \text{Start Time})$ (Units may vary based on team context, capability, cadence; hours, days, weeks, months. May also vary based on calendar time vs. work days. Results with fractional values are rounded up to the next unit.)

- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 - ↓ 4. **Not Useful**
 - ↑ **Impact**
 5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 - ↓ 8. **Not Used**

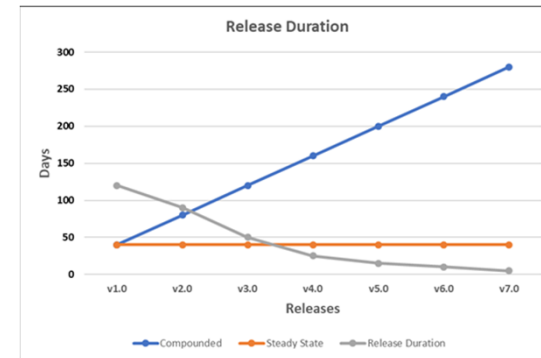
(select 1 from each)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

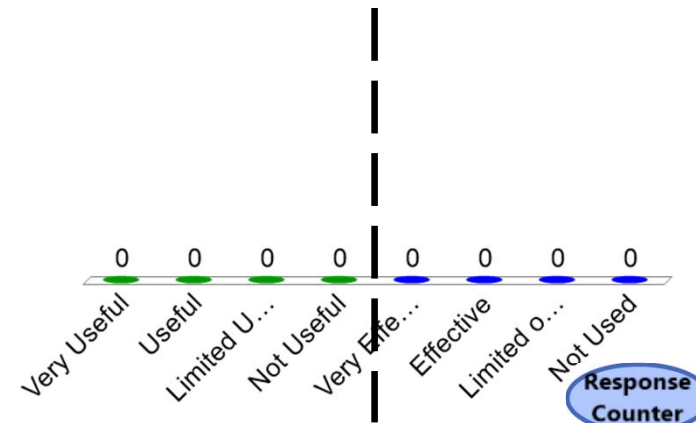
Is **Release Frequency** a **useful** measure, and how **effectively** is it **used** to provide insight and impactful action in **your** organization?

Information Need and Measure Description	
Information Need	How long does it take to develop and release viable products? (<i>duration/time to release new capability</i>) Are release candidates being produced at the cadence needed? (<i>frequency of releases</i>) How long (duration/time) and how much effort/cost(?) does it take to transition candidate products to a completed product baseline release? (<i>duration and effort/cost to deploy release candidates</i>)
Base Measure 1	Start and end dates for a product baseline release (date)
Base Measure 2	Effort hours to transition candidate products to a completed product baseline release
Derived Measure 1	Release duration = (release end date) – (release start date) <ul style="list-style-type: none"> Time to Minimal Viable Product (MVP) = (end date for MVP₁ release) – (start date for MVP₁ release) (<i>initial release of useful capability</i>) Time to Next Viable Product (NVP_n) = (end date for NVP_n release) – (end date for prior NVP_{n-1} release) (<i>subsequent deployments of incremental capability</i>) Time to release a Candidate Product (internal) (e.g., nightly, sprint, increment, other) Time to release an operational product
Derived Measure 2	Average release duration = $\sum (\text{release duration}) / (\# \text{ of releases})$ <i>Note: weighting can be used to emphasize the most recent releases.</i>
Derived Measure 3	Average release transition time = $\sum (\text{release transition time}) / (\# \text{ of releases})$

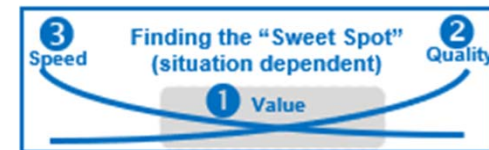


- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 4. **Not Useful**
- ↓ **Impact**
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 8. **Not Used**

(select 1 from each)

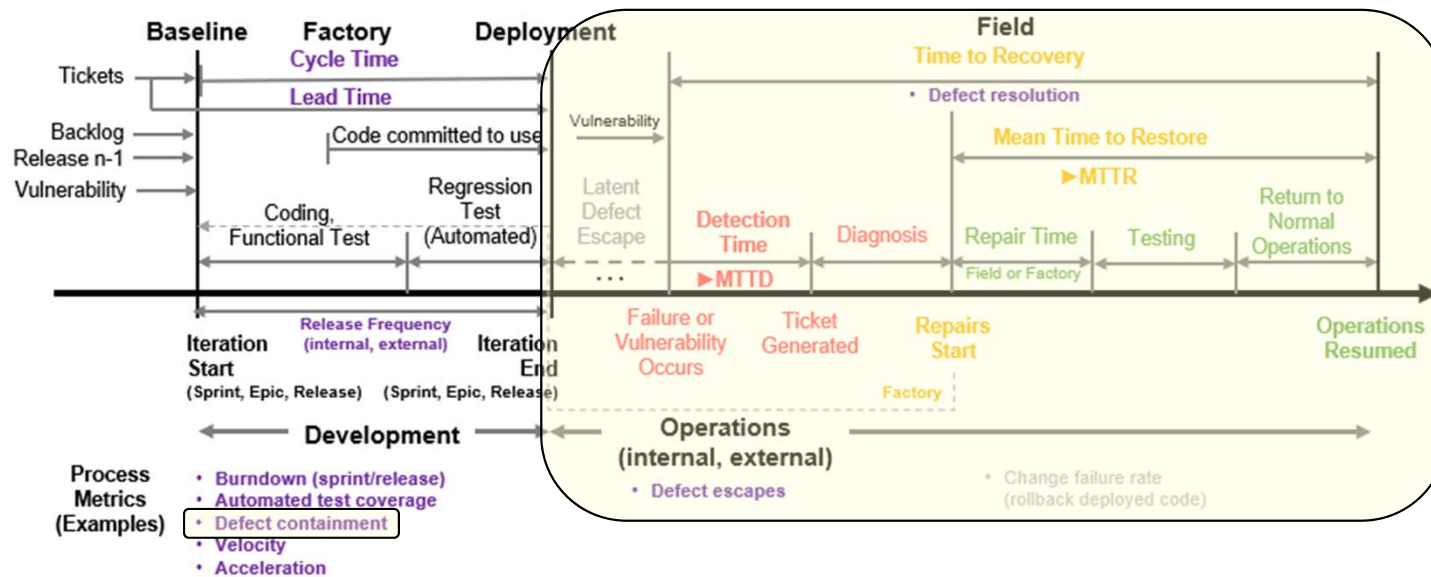


Quality Measures



Concepts:

- Speed can not be optimized without also managing quality
- Quality objectives will vary according to context and domain
- Code quality is integrated into the factory pipeline processes
- Automated verification to the extent practical
- Defect measures are based primarily on escapes from development to operations (internal, external)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

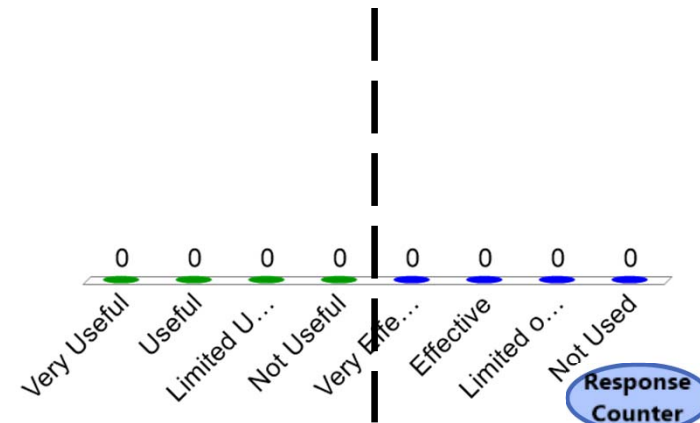
Is **Defect Containment** a **useful** measure, and how **effectively** is it **used** to provide insight and impactful action in **your** organization?

Information Need and Measure Description	
Information Need	<ul style="list-style-type: none"> When were defects introduced into the system (originated) and when were these defects discovered? How effective was the defect discovery process? What was the relative difference between Defect Originated and Defect Discovered? How many defects were not caught until later iterations? Why did these defects escape detection? How can we discover defects earlier in the product lifecycle?
Base Measure 1	Number of defects originated and discovered in same iteration (integer scale)
Base Measure 2	Number of defects discovered one iteration past the iteration originated (integer scale)
Base Measure 3	Number of defects discovered two or more iterations past the iteration originated (integer scale)
Base Measure 4	Number of defects that originated as Legacy defects (e.g., previous projects)
Base Measure 5	Number of defects that have an unknown origin

Defect Containment		Defect Discovered (iteration)							
As of 11 Jun 19		1	2	3	4	5	6		
Defect Originated (iteration)	Unknown	0							
	Legacy	0							
	1	82	29	2	19	17	4	11	
	2	123		27	71	6	7	12	
	3	282			122	60	29	71	Blank 0 %
	4	112				16	2	94	Threshold 41 %
	5	7					5	2	Goal 21 %
6	54						54	Expected 38 %	
	Total	660	29	29	212	99	47	244	

- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 4. **Not Useful**
- ↓ **Impact**
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 8. **Not Used**

(select 1 from each)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

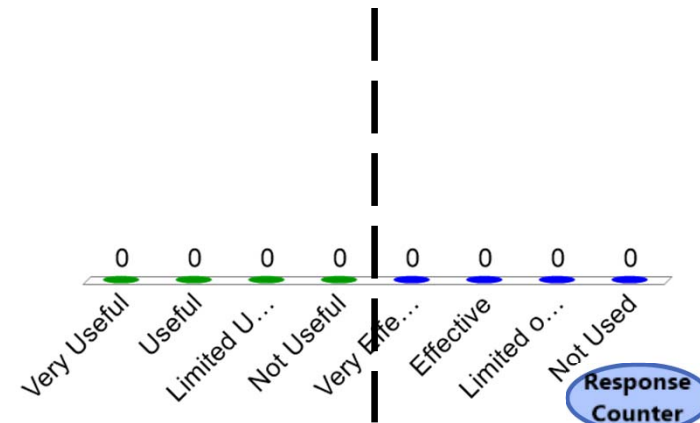
Is **Defect Escapes** a **useful** measure, and how **effectively** is it **used** to provide insight and **impactful** action in **your** organization?

Information Need and Measure Description	
Information Need	For each release (iteration), how many defects were found in internal testing? How many defects were fielded? What is the ratio of fielded defects to all defects? How many defects were found before release to the customer? How many defects were found after release to the customer?
Base Measure 1	Internal Defects (integer scale). Defects found by the development team before release to the internal or external customer. The customer is the team that receives the delivered product. Include priorities 1-3.
Base Measure 2	Fielded Defects (integer scale). Defects found after release to the internal or external customer. Include priorities 1-3.
Derived Measure 1	Defect Escape Ratio = Fielded Defects / (Internal Defects + Fielded Defects)

Release	Defects		Total Defects	Ratio
	Internal	Delivered		
Release 1.0	48	11	59	19%
Release 1.1	55	6	61	10%
Release 1.2	31	4	35	11%
Release 2.0	64	6	70	9%
Release 2.1	55	8	63	13%
Release 2.2	48	4	52	8%
Release 2.3	31	3	34	9%
Release 3.0	20	1	21	5%

- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 4. **Not Useful**
- ↓ **Impact**
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 8. **Not Used**

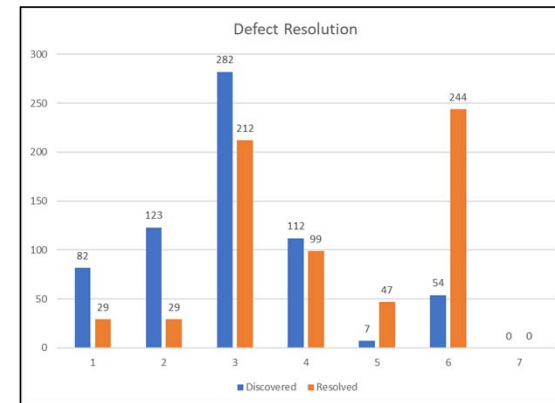
(select 1 from each)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

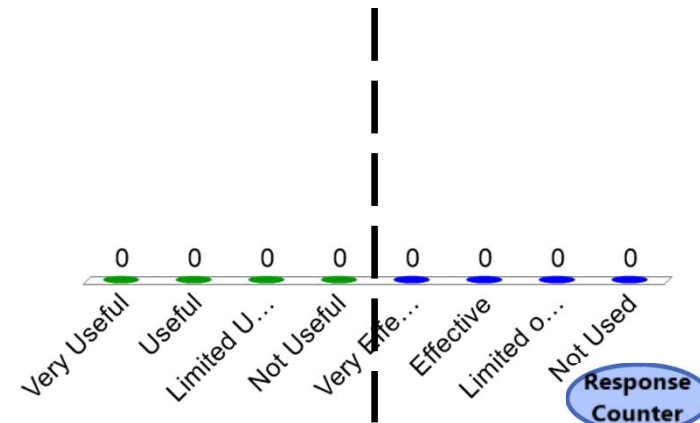
Is **Defect Resolution** a **useful** measure, and how **effectively** is it **used** to provide insight and impactful action in your organization?

Information Need and Measure Description	
Information Need	<ul style="list-style-type: none"> When were discovered defects resolved? How effective was the defect resolution process? How can we resolve defects earlier in the product lifecycle?
Base Measure 1	Number of defects discovered per iteration (integer scale)
Base Measure 2	Number of defects resolved per iteration (integer scale)
Derived Measure 1	Total number of defects discovered in each iteration
Derived Measure 2	Total number of defects resolved in each iteration (integer scale)
Derived Measure 3	Expected Percentage (Resolved) = the number of defects that are resolved in the same iteration they were discovered in (Defect Resolved is the same as Defect Discovered) / the total number of defects
Derived Measure 4	Goal Percentage (Resolved) = the number of defects that are resolved 1 iteration after being discovered / the total number of defects
Derived Measure 5	Threshold Percentage (Removed) = the number of defects that are resolved more than 1 iteration after being discovered / the total number of defects
Derived Measure 6	Defect Age (for active/non-closed defects) = Current Increment – Increment Discovered
Derived Measure 7	Defect Cycle Time (for closed defects) = Increment Resolved – Increment Discovered



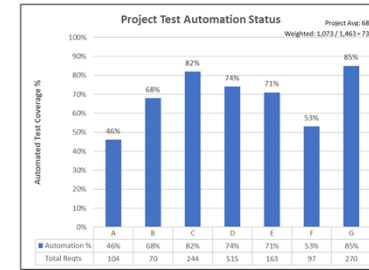
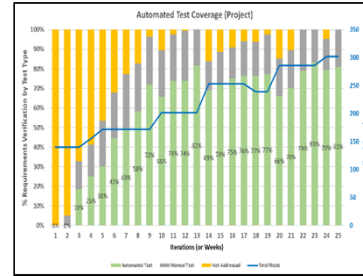
- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
 4. **Not Useful**
- ↓ **Impact**
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 8. **Not Used**

(select 1 from each)



PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

Is **Automated Test Coverage** a **useful** measure, and how **effectively** is it **used** to provide **insight** and **impactful** action in **your** organization?



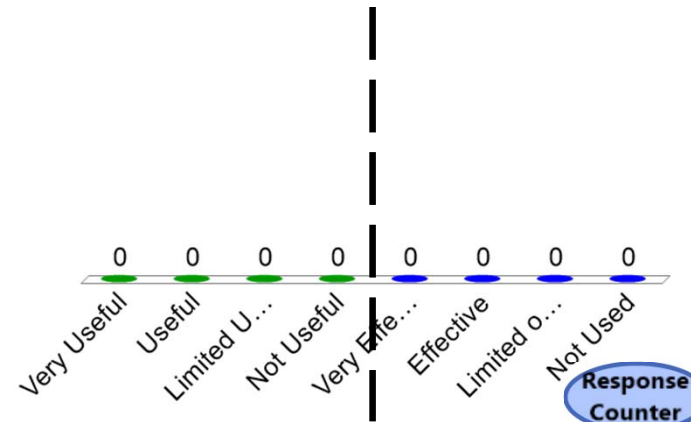
Information Need and Measure Description	
Information Need	How much of the testing is automated? Based on automated test results, what is the quality of the product baseline?
Base Measure 1	Requirements coverage from automated testing (counts, %) Requirements tested by automated test Requirements tested
Base Measure 2	Automated test coverage of code constructs.
Derived Measure 1	% of automated testing coverage for functional requirements
Derived Measure 2	% of automated testing coverage for code constructs (e.g., classes, conditionals, files, lines, packages)

Project

Enterprise

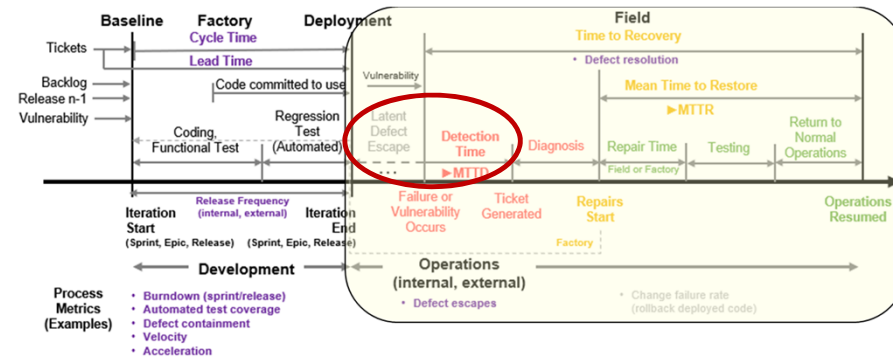
- ↑ **Insight**
- 1. **Very Useful**
- 2. **Useful**
- 3. **Limited Usefulness**
- ↓
- 4. **Not Useful**
- ↑ **Impact**
- 5. **Very Effective**
- 6. **Effective**
- 7. **Limited or No Effectiveness**
- ↓
- 8. **Not Used**

(select 1 from each)



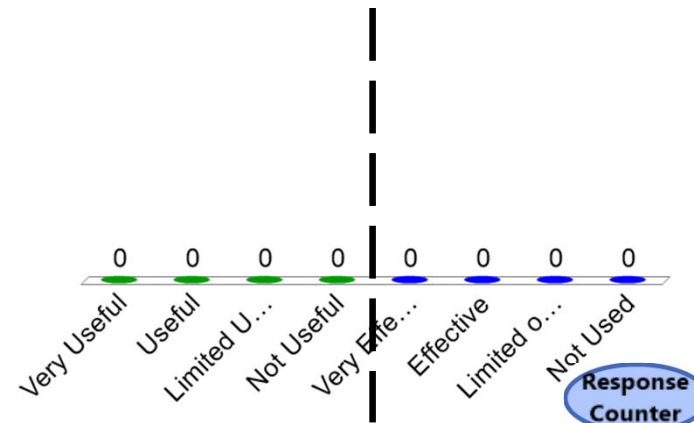
PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

Is **Mean Time to Detect (MTTD)** a **useful** measure, and how **effectively** is it used to provide insight and impactful action in your organization?



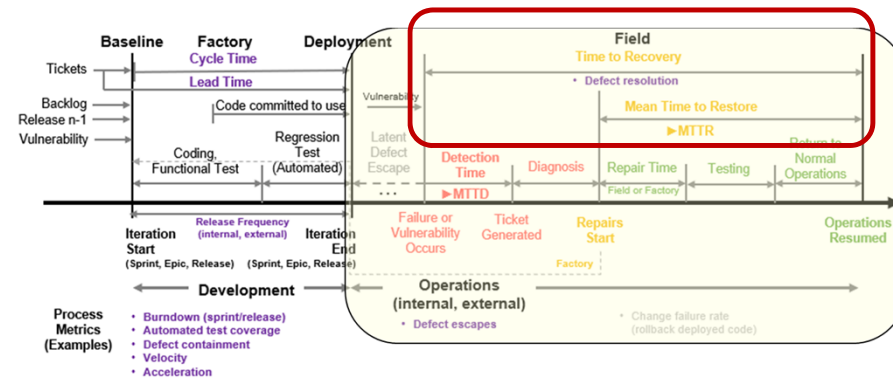
- ↑ **Insight**
- 1. **Very Useful**
- 2. **Useful**
- 3. **Limited Usefulness**
- ↓
- 4. **Not Useful**
- ↑
- 5. **Very Effective**
- 6. **Effective**
- 7. **Limited or No Effectiveness**
- ↓
- 8. **Not Used**

(select 1 from each)



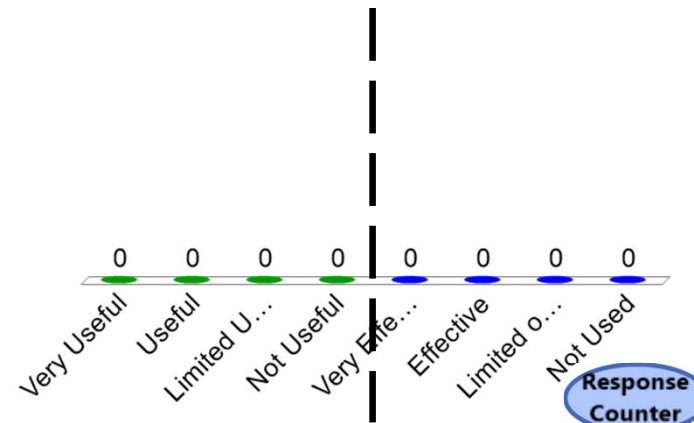
PRACTICAL SOFTWARE AND SYSTEMS MEASUREMENT

Is **Mean Time to Restore (or Recover) (MTTR)** a **useful** measure, and how **effectively** is it used to provide insight and impactful action in **your** organization?



- ↑ **Insight**
1. **Very Useful**
 2. **Useful**
 3. **Limited Usefulness**
- ↓
4. **Not Useful**
- ↑ **Impact**
5. **Very Effective**
 6. **Effective**
 7. **Limited or No Effectiveness**
 8. **Not Used**

(select 1 from each)



Thank you!



Your feedback will help to inform our team progress to validate the consensus PSM measurement framework for continuous iterative development

We plan to publish the initial framework in December 2019

We are seeking motivated volunteers to help further this work – join us!

Cheryl Jones

U.S. Army (FCDD-ACE-QSA)

PSM Project Manager

cheryl.l.jones128.civ@mail.mil

Geoff Draper

L3Harris Technologies

Vice-Chair, NDIA Systems Engineering Division

geoff.draper@l3harris.com

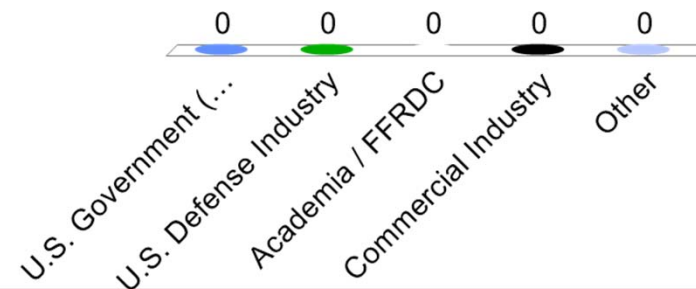
Part I: Evaluation of Information Needs

(deferred due to time constraints)

How would you best characterize your organization?

Response
Counter

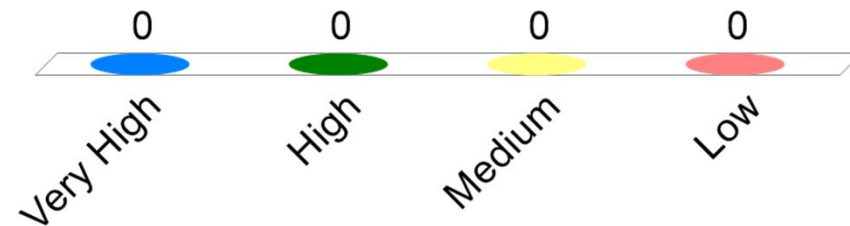
- 1. U.S. Government (DoD, agency)***
- 2. U.S. Defense Industry***
- 3. Academia / FFRDC***
- 4. Commercial Industry***
- 5. Other***



Are new story points, features, or capabilities being delivered as committed?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



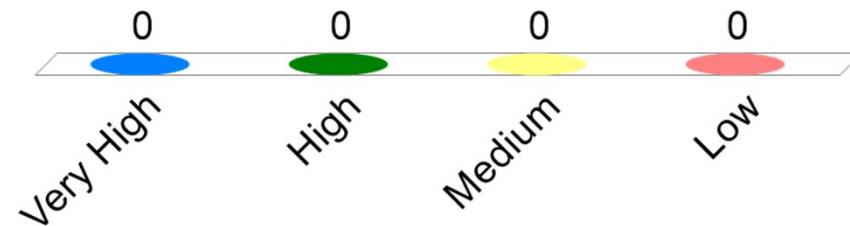
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Schedule and Progress	Work Unit Progress (team, product) Milestone Completion (enterprise)	Are story points delivered as committed?	Are features delivered as committed? Are we still on track to deliver all features per roadmap? (on plan)	Are capabilities delivered as committed?

Response Counter

Do delivered products provide the expected functionality to users when needed?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



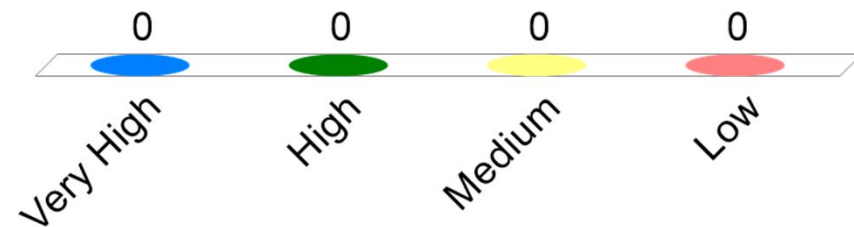
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Schedule and Progress	Work Unit Progress		Did we deliver expected capabilities / features? Is the roadmap still valid?	Is the user satisfied with the delivered products? Do they provide the desired functionality when needed?

Response Counter

How much technical or mission debt exists in the backlog?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



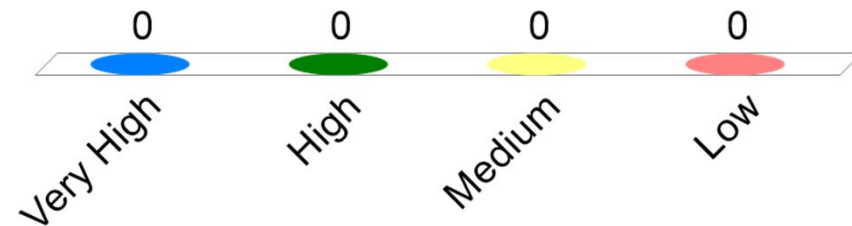
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Schedule and Progress	Work Backlog		How much outstanding technical or mission debt exists?	

Response Counter

Is the product correct? Does new code functionality work as expected?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



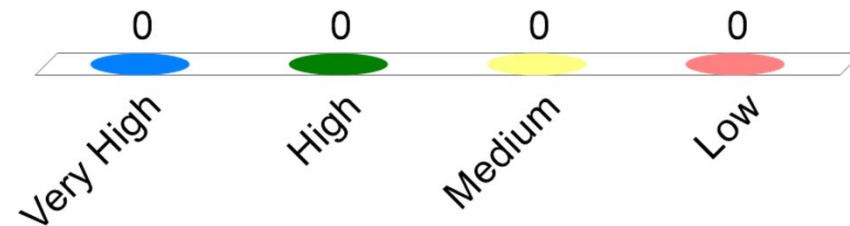
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Product Quality	Functional Correctness	Does new code functionality work as expected?	Does new code functionality work as expected?	Is rework identified and managed?

Response Counter

Does new code break previous functionality? (change failure rate, rollback)

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



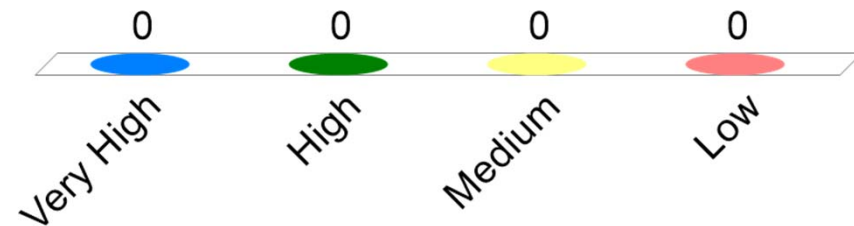
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Product Quality	Functional Correctness	Does new code break previous functionality?	Does new code break previous functionality? (change failure rate, rollback)	

Response Counter

How many defects escape the increment?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



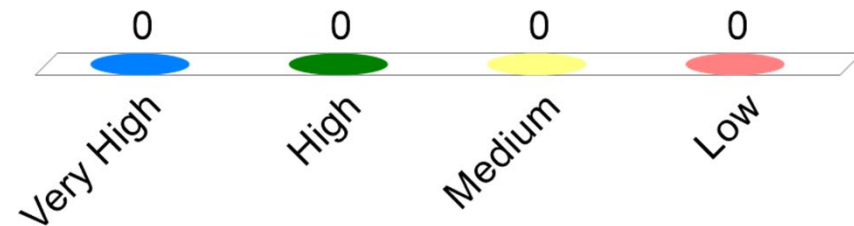
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Product Quality	Functional Correctness		How many defects escape the increment?	

Response Counter

What is the quality of code deployed to the field?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



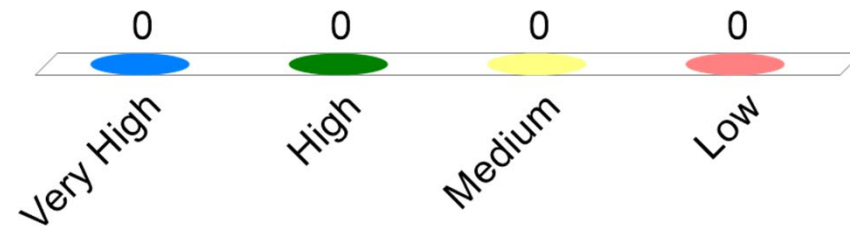
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Product Quality	Functional Correctness		What is the quality of code deployed to the field?	What is the quality of code deployed to the field?

Response Counter

What is the reliability and availability of operational service capabilities?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



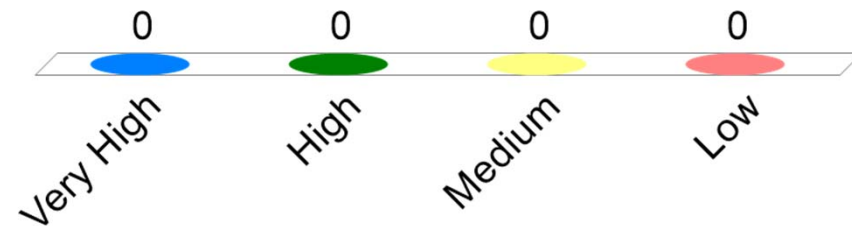
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Product Quality	Supportability - Maintainability Dependability - Reliability		What is the reliability and availability of operational service capabilities?	

Response Counter

How quickly can we address bug reports from the field?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

1. **Very High**
2. **High**
3. **Medium**
4. **Low**



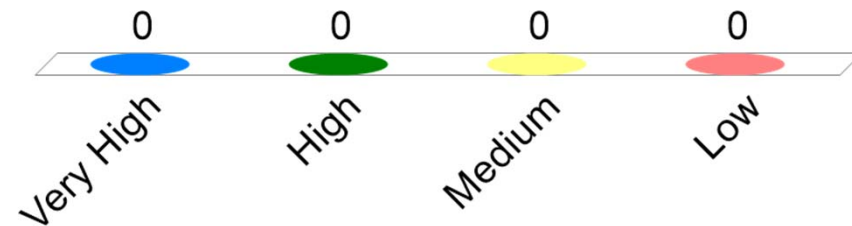
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Process Performance (Process Effectiveness)	Process Efficiency - Speed Supportability - Maintainability Dependability - Reliability		How quickly can we address bug reports from the field?	

Response Counter

Are teams performing as productively as expected?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

1. **Very High**
2. **High**
3. **Medium**
4. **Low**



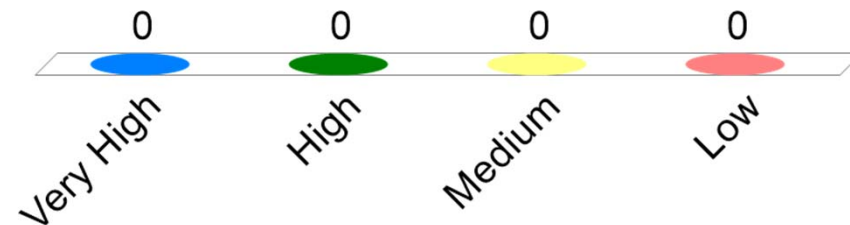
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Process Performance (Process Effectiveness)	Process Efficiency - Speed	Is the team performing as expected?	Are teams performing as expected?	

Response Counter

How long does it take to deploy an identified feature/capability?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



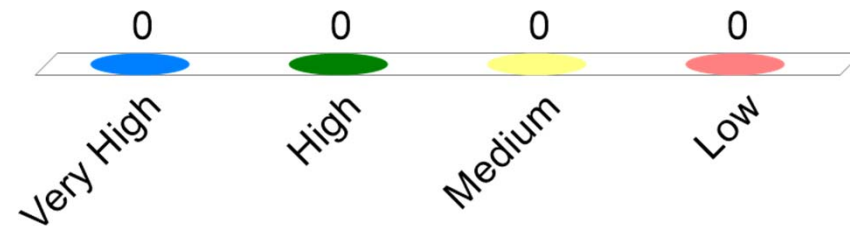
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Process Performance (Process Effectiveness)	Process Efficiency - Speed		How long does it take to deploy an identified feature/capability?	

Response Counter

What is the frequency of product release or deployment?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



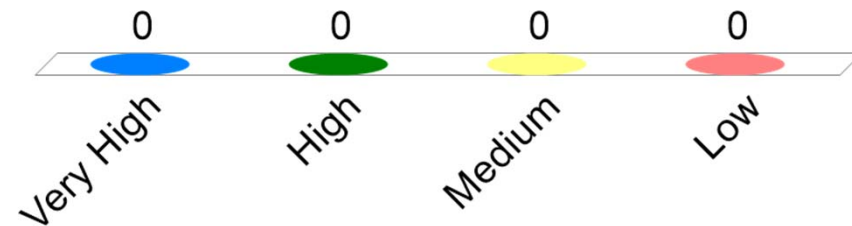
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Process Performance (Process Effectiveness)	Process Efficiency - Speed		What is the frequency of product release or deployment?	What is the frequency of product release or deployment?

Response Counter

How long does it take to release a viable product?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



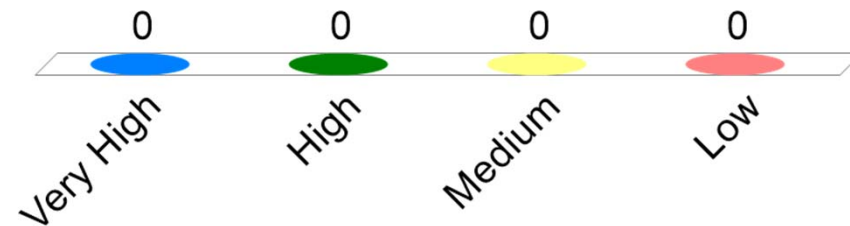
Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Process Performance (Process Effectiveness)	Process Efficiency - Speed	How long does it take to release a viable product?	How long does it take to release a viable product?	How long does it take to release a viable product?

Response Counter

How much of the testing is automated? How often do we perform automated testing?

What importance would you place on this measurement information need for planning and managing continuous iterative SW development?

- 1. Very High**
- 2. High**
- 3. Medium**
- 4. Low**



Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need
Process Performance (Process Effectiveness)	Process Effectiveness		How much of the testing is automated? How often do we perform automated testing?	How much of the testing is automated? How often do we perform automated testing?

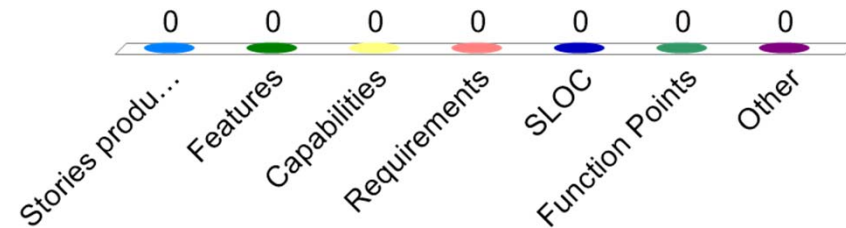
Response Counter

How big is our system?

How should we count size for continuous iterative development programs (e.g., for estimating)?

(pick up to 2 choices)

1. **Stories produced (team)**
2. **Features**
3. **Capabilities**
4. **Requirements**
5. **SLOC**
6. **Function Points**
7. **Other**

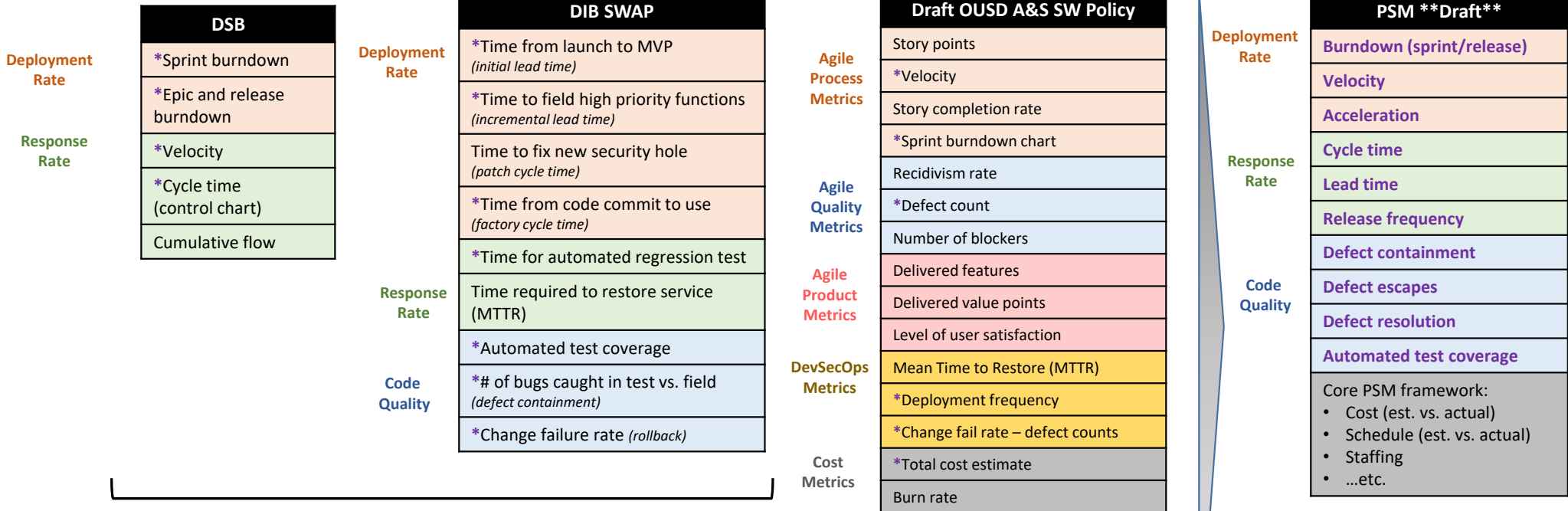


Information Categories	Measurable Concept	Team Information Need	Product Information Need	Enterprise Information Need	Potential Measures
Size and Stability	Functional Size and Stability Physical Size and Stability	How big is our system?	How big is our system?	How big is our system?	Stories produced (team) Features Capabilites Requirements SLOC

Response Counter

BACKUP

Candidate Measures



Industry Survey Feedback (usefulness, effectiveness)

Category	Evaluation and Ranking of CIM Measures					Evaluation and Ranking of CIM Measures					Overall Avg
	1	2	3	4	5	1	2	3	4	5	
Usefulness	0.8	0.7	0.6	0.5	0.4	0.8	0.7	0.6	0.5	0.4	0.7
Sample Count	10	10	10	10	10	10	10	10	10	10	10
Mean	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
Med	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
Std Dev	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Var	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Effectiveness	0.8	0.7	0.6	0.5	0.4	0.8	0.7	0.6	0.5	0.4	0.7
Sample Count	10	10	10	10	10	10	10	10	10	10	10
Mean	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
Med	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
Std Dev	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Var	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25

* = addressed in draft PSM framework

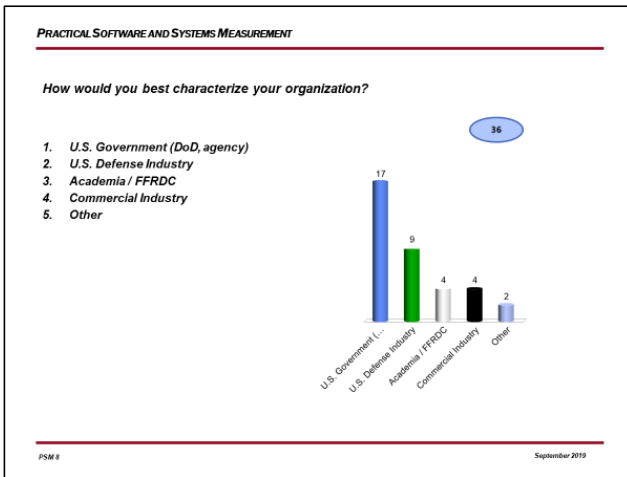
- See PSM framework for details.
- Information categories
 - Measurable concepts
 - Information needs
 - Cross-reference mappings

Additional candidate measures are defined in draft ICM table but not implemented in first release.

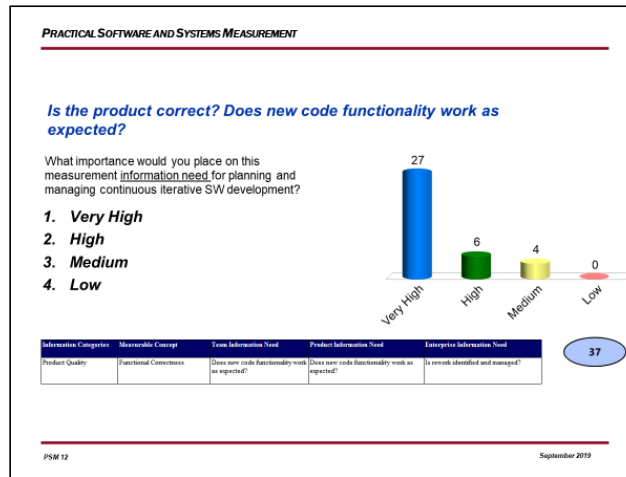
Evaluation of draft PSM agile measurement framework

20th PSM Users' Group Workshop, Sep 15-18, 2019

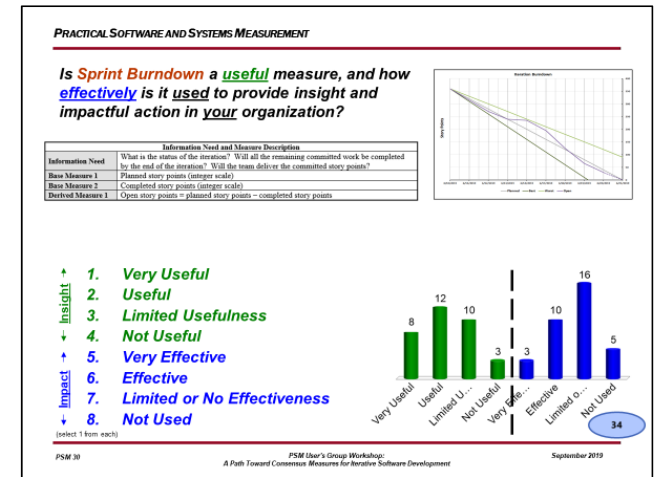
Demographics



Information Needs



Potential Measures



See PSM workshop outputs for details and analyses.
<http://www.psmc.com/UG2019/Workshops/w01.zip>