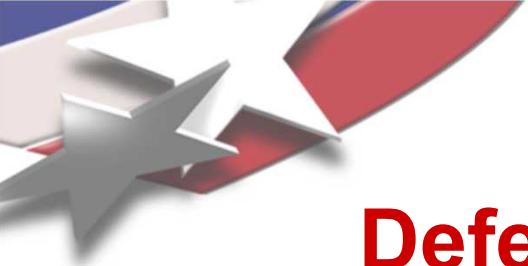


Defect Collection & Analysis – The Basis of Software Quality Improvement

**Joe Schofield,
Sandia National Laboratories
Albuquerque, N. M.**

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy under contract DE-AC04-94AL85000.





Defect Collection & Analysis – The Basis of Software Quality Improvement

Joe Schofield, yhs

Note: The CMMI and SW-CMM are registered copyrights of the Software Engineering Institute

The opinions expressed in the presentation are solely / sorely, the presenter's.

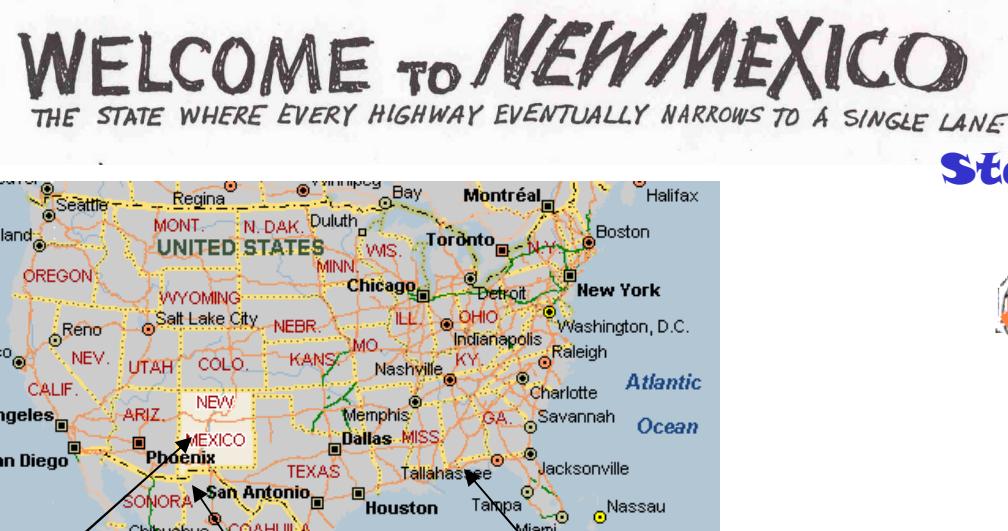
(Geographical) Context Diagram



State mineral



State animal



New Mexico

Mexico

You are here!

State flag & statue



State motto



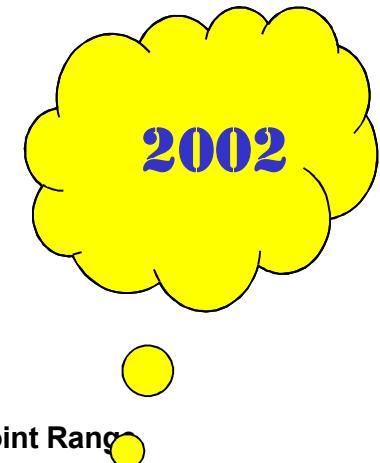


A Quick Look Back & Update on Recent IFPUG / ISMA Presentations

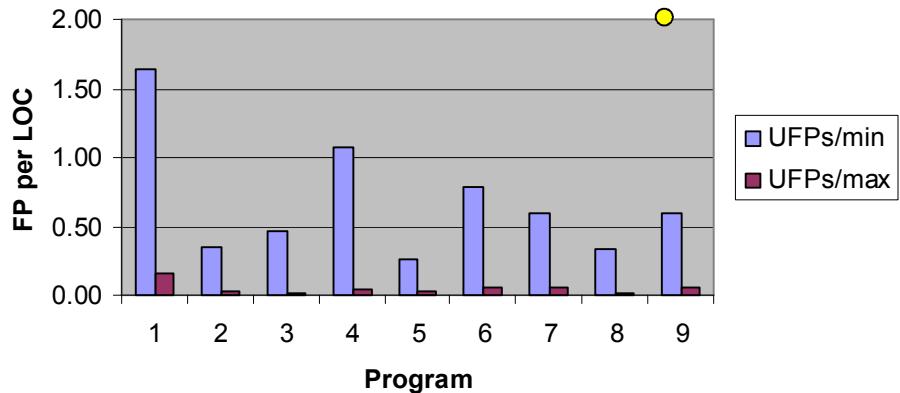
2002	<p>Counting KLOCs – Software Measurement's Ultimate Futility (I can't do this anymore, or who am I fooling?, or why not count ants?)</p> <p><i>The Statistically Unreliable Nature of Lines of Code; CrossTalk, April 2005</i></p> <p><i>A Practical, Statistical, and Criminal Look at the Use of Lines of Code as a Software Sizing Measure ; N.M. SPIN; March, 2004</i></p>
2003	<p>Amplified Lessons from the Ant Hill –</p> <p>What Ants and Software Engineers Have in Common</p> <p><i>Lessons from the Ant Hill - What Ants and Software Engineers Have in Common; Information Systems Management, Winter 2003</i></p>
2004	<p>Applying Lean Six Sigma to Software Engineering</p> <p><i>When Did Six Sigma Stop Being a Statistical Measure?; CrossTalk, April 2006</i></p> <p><i>Lean Six Sigma - Real Stories from Real Practitioners; Albuquerque, N.M.; N.M. SPIN; August 2005</i></p> <p><i>Six Sigma & Software Engineering: Complement or Collision; Albuquerque, N.M.; N.M. SPIN; August, 2004</i></p>
2005	<p>Defect Collection & Analysis – The Basis of Software Quality Improvement</p> <p><i>Defect Management through the Personal Software Process(SM); CrossTalk, September 2003</i></p> <p><i>The Team Software ProcessSM - Experiences from the Front Line; Software Quality Forum; Arlington, Virginia, March; 2003</i></p> <p><i>Measuring Software Process Improvement - How to Avoid the Orange Barrels; System Development, December 2001</i></p> <p><i>Usable Metrics for Software Improvement within the CMM; Software Quality Forum 2000; Santa Fe, N.M.; April, 2000</i></p>

Min and max values for “C” code compared to Function Point size over 9 programs (n = 49)

	*P1	P2	P3	P4	P5	P6	P7	P8	P9
<i>Min</i>	22	20	15	13	27	23	25	21	25
<i>Max</i>	221	311	336	289	270	306	242	383	284
<i>UFPs</i>	36	7	7	14	7	18	15	7	15
<i>UFPs/min</i>	1.64	0.35	0.47	1.08	0.26	0.78	0.60	0.33	0.60
<i>UFPs/max</i>	0.16	0.02	0.02	0.05	0.03	0.06	0.06	0.02	0.05
<i>Variance</i>									
<i>Range</i>	10.05	15.55	22.40	22.23	10.00	13.30	9.68	18.24	11.36
									14.76



"C" Min/Max Function Point Range



Largest min to max variance is > 22, smallest is almost 10, average is almost 15.

Note that in these three examples, variance and averages increased as the population increased.

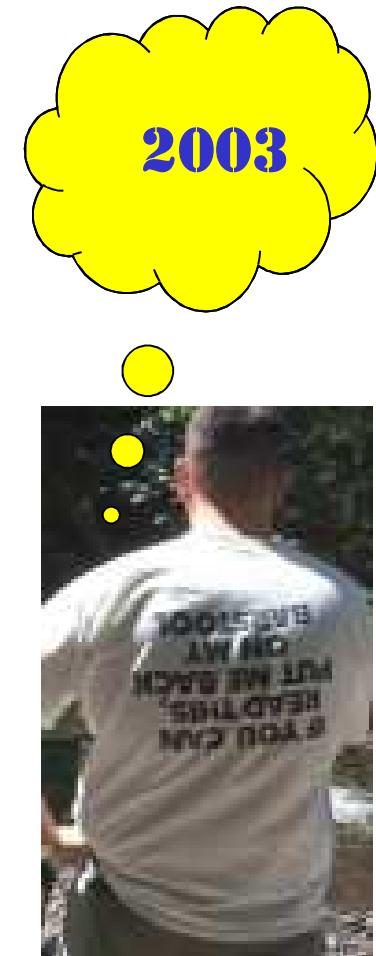


Software engineers are smarter than ants, right?

Observation: When ants underestimate the size of a job, they compensate with waves of more ants. Most software projects cannot afford this tactic.

Lesson: Use reliable sizing measures like Function Points to assess progress. Avoid the practice of counting lines of code as a measure of size or progress.

Reference: *A Practical, Statistical, and Criminal Look at the Use of Lines of Code as a Software Sizing Measure*, Schofield, Structured Development Forum, March, 2003



When Lean Six Sigma Isn't (cont'd)

“What if” the *sigma shift* went to the right – a teraflop example

TeraFlops machine

1T floating point operations instructions per second =

3 defects per 100 seconds =

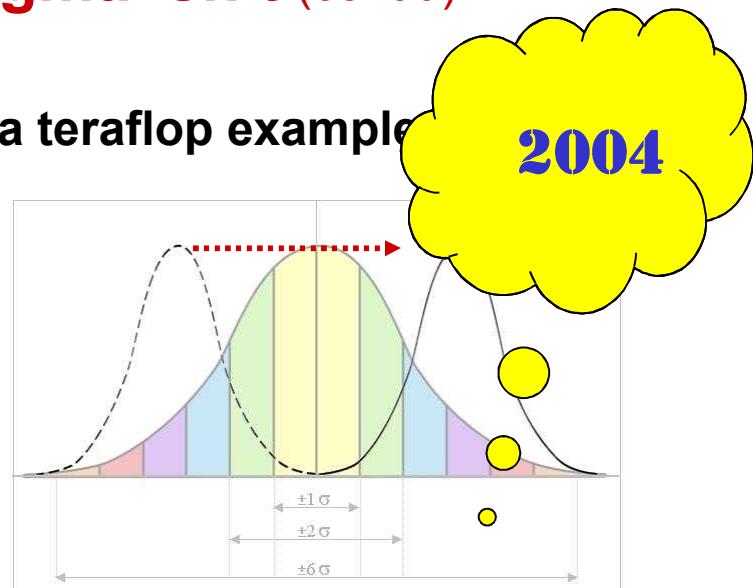
108 defects per hour =

18,144 per week =

943,488 DEFECTS per year =

50M+ a year at “shifted 6 sigma” (4.5 sigma)

(these numbers are rounded down)



¹ PetaFlops machine

predicted to be ready by 2005 or 2006

1,000 times faster than a 1TFlop machine =

943,488,000 defects per year @ 7.5 sigma =

50B (that's BILLION) at “shifted 6 sigma”

Who can repair / afford / manage that many defects?

PETAFLOP Imperative; Informationweek; June 21, 2004; pgs. 55 – 62

² IBM's Gene/L at Lawrence Livermore National Lab operates @ 70.72TF

IBM will increase the speed to 360 TF in 2005

U.S. Regains Top Supercomputer Spots; Informationweek; November 15, 2004; pg. 28



(Back to Defects . . .)

The Business Case for Defect Removal

Software defects cost the U.S. \$59.6B a year¹

38 percent of polled organizations have no SQA program²

Software technicians in Panama are charged with second degree murder after 27 patients received overdoses of gamma rays; 21 have died in 40 months³

BMW, DaimlerChrysler, Mitsubishi, and Volvo experience product malfunctions (engine stalls, gauges not illuminated, wiping intervals, wrong transmission gears) due to software⁴

A 2002 GAP report showed that spreadsheet errors at NASA contributed to a \$644M misstatement in 1999⁵

SPAM will cost the world \$50B in lost productivity and other expenses . . . according to Ferri Research⁶

Medical staff report 38 percent defect rate while using computerized physician order entry (CPOE) systems in determining low dose for infrequently used medications.⁷

The FBI's \$170M virtual case file project went through 10 program managers before being cancelled.⁸

¹ Informationweek, *Behind the Numbers*, March 29, 2004; pg 94

² CIO, *By the Numbers*, December 1, 2003, pg 28

³ Baseline – The Project Management Center, *We Did Nothing Wrong*, March 4, 2004

⁴ Informationweek, *Software Quality*, March 15, 2004; pg 56

⁵ CIO, *Essential Technology*, May 15, 2005; pg 74

⁶ Informationweek, February 28, 2005; pg 18

⁷ CIO, *Medication Systems*, June 1, 2005; pg 28

⁸ CIO, *Why the G-Men Aren't IT Men*, June 15, 2005; pg 44

The Business Case for Defect Removal (continued)

December Data Exposures



JUSTICE Social Security numbers and other personal data of an unknown number of people involved in various cases can be seen on its Web site



MARRIOTT Social Security and credit-card numbers and other personal data of 206,000 employees and customers went missing



FORD A computer with personal data, including Social Security numbers, on 70,000 current and former employees was stolen



ABN AMRO A tape containing data on about 2 million customers was lost as it was being transported; the tape was later found



SAM'S CLUB At least 600 customers who bought gas at Sam's Club stores between Sept. 21 and Oct. 2 had credit-card data stolen

Informationweek; 1/2/2006; pg. 19

Room for Improvement

ALIGNMENT In a recent survey of more than 1,400 midsize organizations in the United States, Canada and the United Kingdom conducted by Info-Tech Research Group, 95 percent of respondents said their IT departments are failing to deliver projects either on time or that meet the expectations of business executives. It should hardly come as a surprise, therefore, that only 11 percent of the organizations that responded view technology as a "strategic weapon."

"The fact that almost every IT department is failing on some of their projects is both a major contributor to, and a major symptom of, the misalignment of business and IT," says Frank Koelsch, executive vice president for corporate strategy and research at the Info-Tech Research Group.

CIO; 5/1/2005; pg. 28

According to "IT Priorities 2005," a survey by Info-Tech Research Group, organizations cited these top three reasons that projects fail.

Unrealistic time frames	68%
Insufficient staffing	64%
Poorly defined project scope	62%

(Organizations could select more than one answer.)

Terror Data Collection System a 'Boondoggle'

BY SIOBHAN GORMAN

The Baltimore Sun

WASHINGTON—A program that was supposed to help the National Security Agency identify electronic data crucial to the nation's safety is not up and running more than six years and \$1.2 billion after it was launched, according to current and former government officials.

The classified project, code-named Trailblazer, was promoted as the NSA's state-of-the art tool for sifting through an ocean of modern day digital communications and uncovering key nuggets to protect the nation against an ever-changing collection of enemies.

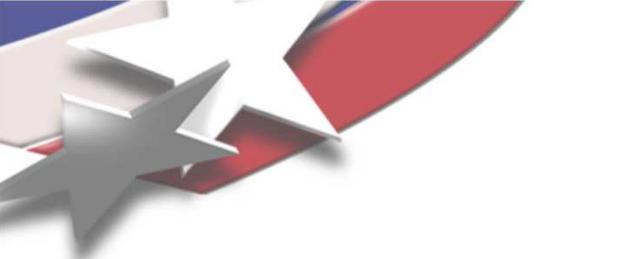
Its main goal when it was launched in 1999 was to allow NSA analysts to connect the 2 million bits of data the agency collects every hour — a task that has grown increasingly complex with the advent of the Internet, cell phones and instant messaging — and enable them quickly to identify the

most important information.

A major failure leading up to Sept. 11, 2001 involved communications intelligence, investigators found. More than 30 hints of the impending attack had been collected in the previous three years but had gone unnoticed in the NSA's databases, according to a joint congressional inquiry into pre-Sept. 11 intelligence operations.

The NSA initiative, which was designed to spot and analyze such information, has resulted in little more than detailed schematic drawings filling almost an entire wall, according to intelligence experts. After an estimated \$1.2 billion in development costs, only a few isolated analytical and technical tools have been produced, said an intelligence expert with extensive knowledge of the program.

Trailblazer is "the biggest boondoggle going on now in the intelligence community," said Matthew Aid, who has advised three federal commissions and panels that investigated the Sept. 11 intelligence failures.



The Business Case for Defect Removal (continued)

GM Recalls 900,000 Chevy, GMC Pickups

Tailgate Cable Might Break

BY KEN THOMAS
The Associated Press

WASHINGTON — General Motors Corp. said Friday it is recalling about 900,000 pickup trucks worldwide to fix problems with tailgate cables that can corrode and break when loads are placed on them.

The recall involves 1999-2000 models of the Chevrolet Silverado and GMC Sierra trucks. GM said there have been 84 injuries, most of them minor scrapes and bumps, but no crashes or deaths linked to the problem.

GM, the world's No. 1 automaker, said the galvanized, braided-steel support cables that keep the tailgates in place can corrode or fray over time, "which can make them break," Adler said.

Albuquerque Journal; 3/2006; Business page

time because of moisture seeping through cracks in the plastic sheathing of the cable or entering between the cable's metal strands.

The automaker had recalled about 4 million 2000-2004 pickups worldwide in March 2004 because the tailgates could break without warning. The recall involved a broader range of vehicles, including the Chevrolet Silverado, GMC Sierra, Chevrolet Avalanche and Cadillac Escalade EXT trucks.

GM spokesman Alan Adler said the vehicles covered by the new recall had different materials used in the support cables and involved fewer complaints. He said the automaker decided to issue the recall after "an exhaustive investigation."

"We worked with (the National Highway Traffic Safety Administration) and decided to get this house in line," Adler said.

Informationweek; 6/5/2006; pg. 18

House Wrecks Valued at \$400M

The Associated Press

VALPARAISO, Ind. — A house erroneously valued at \$400 million is being blamed for budget shortfalls and possible layoffs in municipalities and school districts in northwest Indiana.

An outside user of Porter County's computer system may have triggered the mess by accidentally changing the value of the Valparaiso house. The house had been valued at \$121,900 before the glitch.

County Treasurer Jim Murphy said the home usually carried about \$1,500 in property taxes; this year, it was billed \$8 million.

U.K. HEALTH SYSTEM

One Sick IT Project

CALLED "THE GREATEST IT DISASTER IN HISTORY" one London columnist, the U.K. National Health Services' program to create nationwide e-health records and upgrade the medical system's IT infrastructure is in trouble. Britain's Labor government says the program will end up costing more than \$55 billion — a whopping \$26 billion over budget.

The project, run by Accenture, Computer Sciences Corp., Fujitsu, and others, has been plagued by software problems and resistance from physicians. The vendors face big penalties if deadlines aren't met. In March, Accenture took a \$450 million hit as a result. It's enough to make you sick.

—PAUL McDougall

Informationweek; 6/5/2006; pg. 18



Basic Measures for Defects

Defect – a product anomaly such as an omission or imperfection; faults in software sufficiently mature for test or operation. Adapted from the IEEE 982.1

Defect Attributes:

Defect injection

Defect detection

Defect removal cost

Defect type

Defect status

Defect severity

Peer Review times

Defect repair cost

Derivable Defect Measures:

Defect per Function Point

Average cost of defect by phase

Average cost of defect by work product

Defect leakage

Predicted number of defects (historic DFP)

Cost of Peer Reviews

Cost of defect removal

Cost of testing

ROI for defect removal (Cost of prevention and removal vs. cost of repair)



Applying DMAIC (Six Sigma) to Defect Data

Actual cost benefit figures for software development

Measures:

Define (opportunities)

An Organization Definition

Measure (performance)

Peer Reviews & Defects

Analyze (opportunity)

Defect List & Analysis

Improve (performance)

Process Focus & Change

Control (performance)

Sustained Measurement & Improvement

Whats

Hows



Applying DMAIC to Defect Data (cont'd)

Required items are **bold**.

Attribute	Value
Discovered By	Change Request
Peer Review Item	
Detection Phase	Planning
Injection Phase	Planning
Defect Type	Completeness
Defect Severity	Aesthetic
Cost to Repair	
Description/Class	
Disposition	

Measure / Record

Analyze

Defect Types

Category	Count
Failure	~20
Functional	~165
Aesthetic	~125

Injected Defects for 12 Projects

Phase	Defects
Ops.	~115
Deploy.	~40
Impl.	~75
Design	~95
Analysis	~95
Planning	~40

Distribution of Defect Cost to Repair

Cost Range	Count
Cost not recorded	~40
More than \$100	~5
Less than or equal to \$100	~50
Less than or equal to \$20	~35
Less than or equal to \$5	~145

Applying DMAIC to Defect Data (cont'd)

[Modify or Input Defects](#)

SILC Review Phase	Discovered By	Defect Type	Total Defects	Total Cost	Cost Per Defect
Planning	Change Request	Completeness	20	707	35.35
Planning	Peer Review	Completeness	91	3867	42.49
Planning	Peer Review	Consistency	21	667	31.76
Planning	Peer Review	Corrective	33	1481	44.88
Planning	Test Plan	Completeness	1	4	4.00
Analysis	Change Request	Completeness	6	180	30.00
Analysis	Change Request	Corrective	6	60	10.00
Analysis	Peer Review	Completeness	124	2900	23.39
Analysis	Peer Review	Consistency	103	1968	19.11
Analysis	Peer Review	Corrective	109	1890	17.34
Design	Change Request	Completeness	3	160	53.33
Design	Change Request	Corrective	4	170	42.50
Design	Peer Review	Completeness	265	7406	27.95
Design	Peer Review	Consistency	59	1313	22.25
Design	Peer Review	Corrective	162	2054	12.68
Design	Test Plan	Completeness	2	8	4.00
Design	Test Plan	Consistency	1	6	6.00
Design	Test Plan	Corrective	4	124	31.00
Implementation	Change Request	Completeness	2	80	40.00
Implementation	Change Request	Corrective	8	1337	167.13
Implementation	Peer Review	Completeness	63	2125	33.73
Implementation	Peer Review	Consistency	55	1909	34.71
Implementation	Peer Review	Corrective	76	2572	33.84
Implementation	Test Plan	Completeness	36	3801	105.58
Implementation	Test Plan	Consistency	15	1146	76.40
Implementation	Test Plan	Corrective	85	3151	37.07
Deployment	Change Request	Corrective	4	67	16.75
Deployment	Peer Review	Completeness	29	200	6.90
Deployment	Peer Review	Consistency	1	4	4.00
Deployment	Peer Review	Corrective	7	38	5.43
Operational	Change Request	Completeness	5	408	81.60
Operational	Change Request	Consistency	4	195	48.75
Operational	Change Request	Corrective	12	1215	101.25
Operational	Test Plan	Corrective	11	1319	119.91

Find
Remove
Prevent

Applying DMAIC to Defect Data (cont'd)

Find
Remove
Prevent

Artifact Reviewed	Total Defects	Total Cost	Cost Per Defect
External Interfaces Definition	94	1612	17.15
Information Model	57	525	9.21
Internal Components Definition	67	1507	22.49
Other: Business Rules	5	13	2.60
Other: Business Rules and Use Cases	33	122	3.70
Other: CSA Administrator Documentation	2	20	10.00
Other: CSA User Documentation	2	20	10.00
Other: EP Interim Solution: Source Code	13	308	23.69
Other: External Interfaces Definition - Admin	5	9	1.80
Other: External Interfaces Definition - Report	9	29	3.22
Other: External Interfaces Definition - Wizard	4	47	11.75
Other: Information Model, Data Dictionary	7	22	3.14
Other: Interim Solution Test Plan	6	159	26.50
Other: Internal Components Definition--Admin	1	1	1.00
Other: Internal Components Definition--Course	3	266	44.33
Other: RS2 User Process Model	5	70	14.00
Other: RS3 SW Requirements & Design Specification	1	50	50.00
Other: RS3 Software Source Code & Executables	5	610	122.00
Other: RS3 User Process Model	8	165	20.63
Other: RS5 User Process Model	4	14	3.50
Other: RS6 User Process Model	7	300	42.86
Other: RS7 Design	7	200	28.57
Other: RS7 Software Source Code & Executables	3	40	13.33
Other: RS7 User Process Model	8	135	16.88
Other: Software Requirements Specification	27	470	17.41
Other: Test Plan, Information Model, External Interfaces	25	134	5.36
Other: Use Case Diagrams & Textual Use Cases	1	2	2.00
Other: Use Case Model	3	7	2.33
Other: User Documentation	17	101	5.94
Project Plan	129	4005	31.05
Software Source Code & Executables	139	4561	32.81
Test Plan	209	3774	18.06
User Process Model	132	5641	42.73

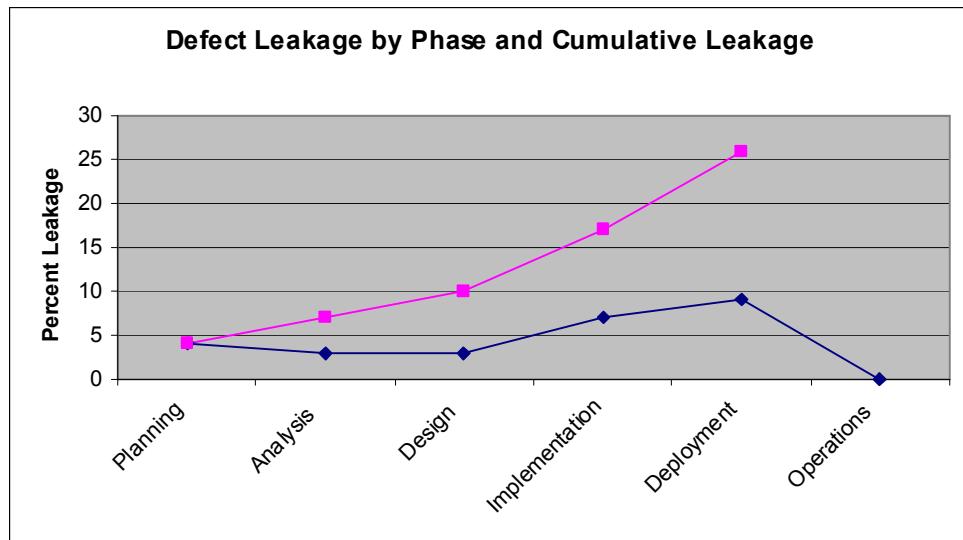
Defect summary by work product

For defect removal, Tom Glib reports some inspection efficiencies as high as 88 percent. Jones, *Software Quality*, pg 215

Applying DMAIC to Defect Data (cont'd)

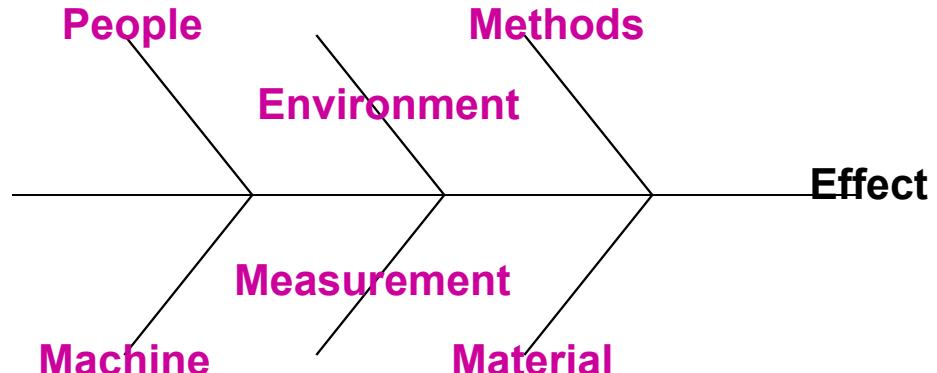
Given:

- 1 Peer Review is performed in Planning
- 2 Peer Reviews are performed in Analysis
- 3 Peer Reviews are performed in Design



Look at Planning & Analysis

- 1) How are so many defects removed in Implementation?
- 2) Does the organization need more Peer Reviews in Planning & Analysis?
- 3) How effective are Design Peer Reviews?



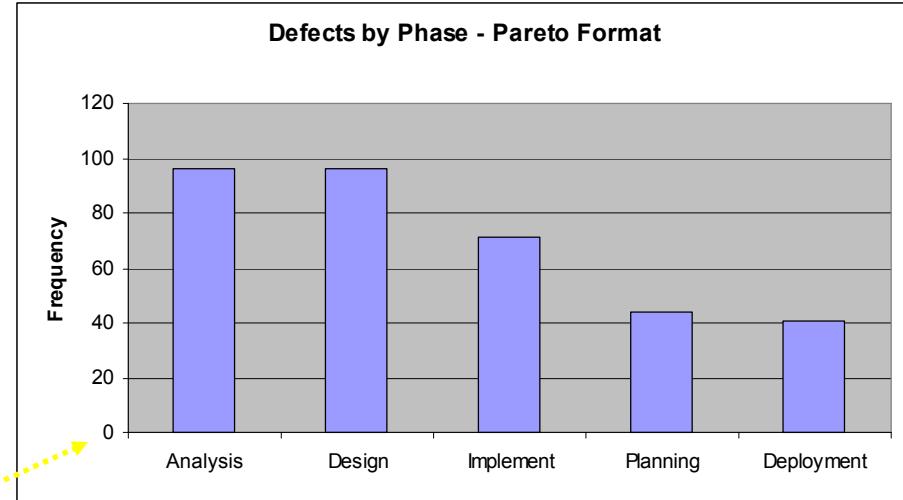
Using Pareto Analysis and Histograms for Defect Data

Required items are **bolded**.

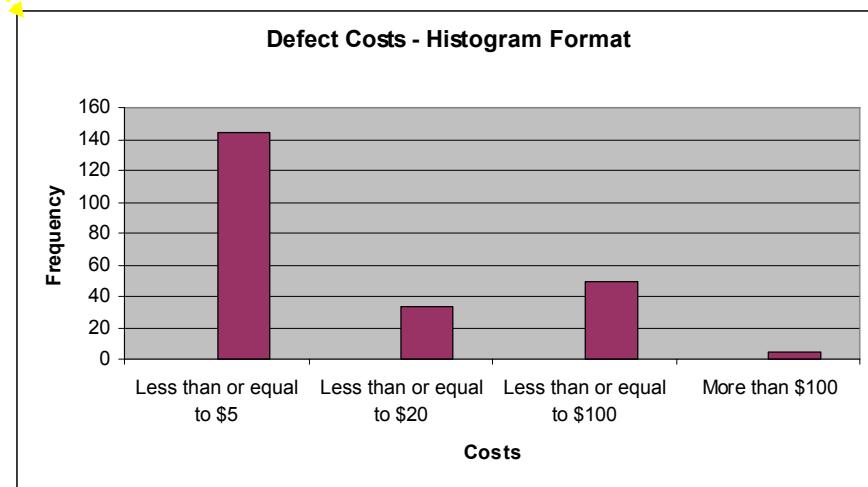
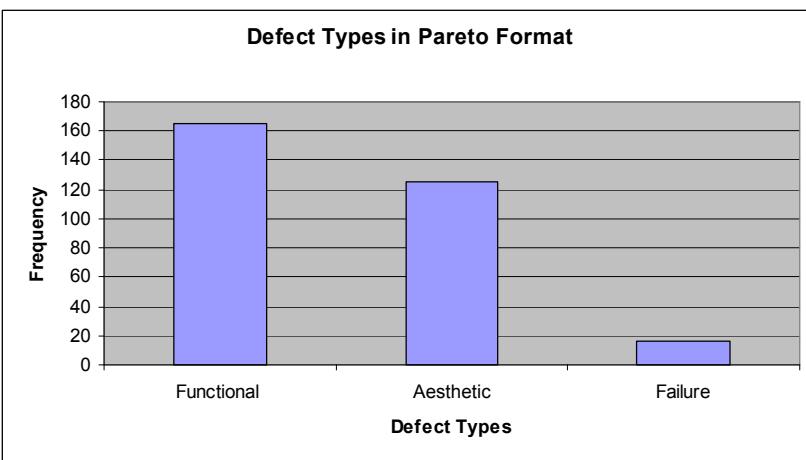
Attribute	Value	
Discovered By	Change Request <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="button" value="..."/>	Peer Review Item <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="button" value="..."/>
Detection Phase	Planning <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="button" value="..."/>	
Injection Phase	Planning <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="button" value="..."/>	
Defect Type	Completeness <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="button" value="..."/>	
Defect Severity	Aesthetic <input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="button" value="..."/>	
Cost to Repair	<input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="text"/>	
Description/Class	<input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="text"/>	
Disposition	<input style="width: 100px; height: 20px; border: 1px solid #ccc; border-radius: 5px; padding: 2px 10px;" type="text"/>	

Buttons: Submit, Reset.

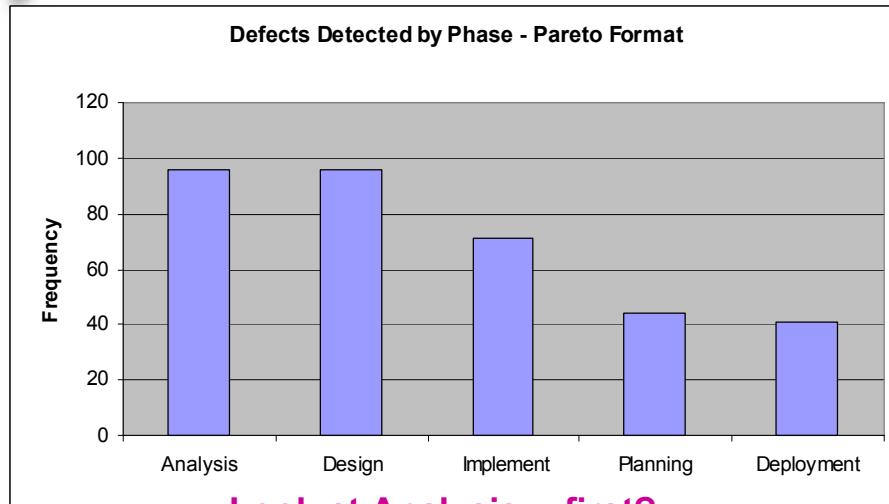
Captured



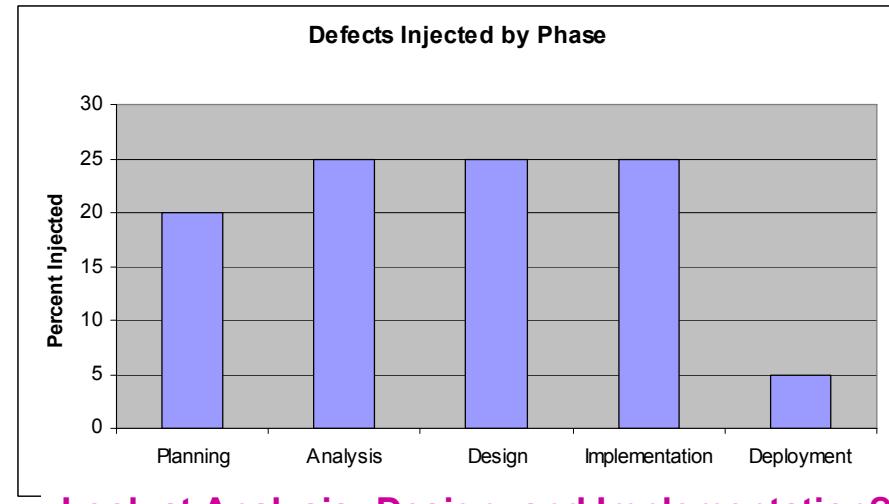
Analyze (Derived)



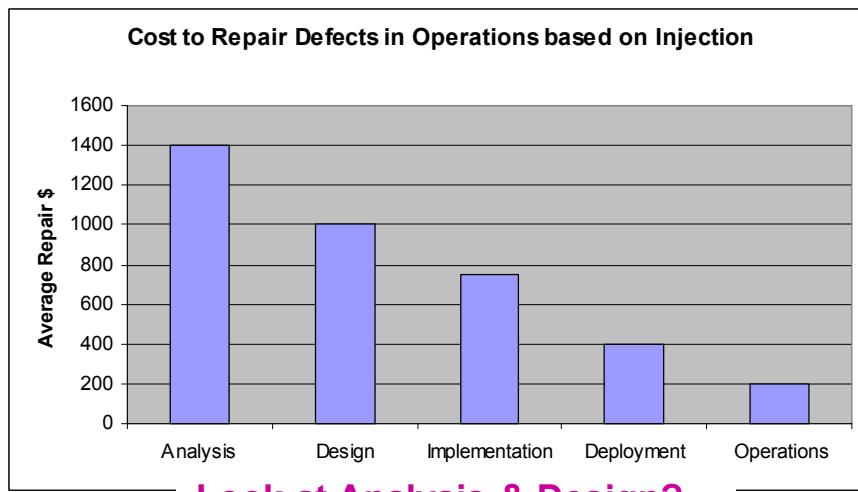
Even Tools Require Thinking



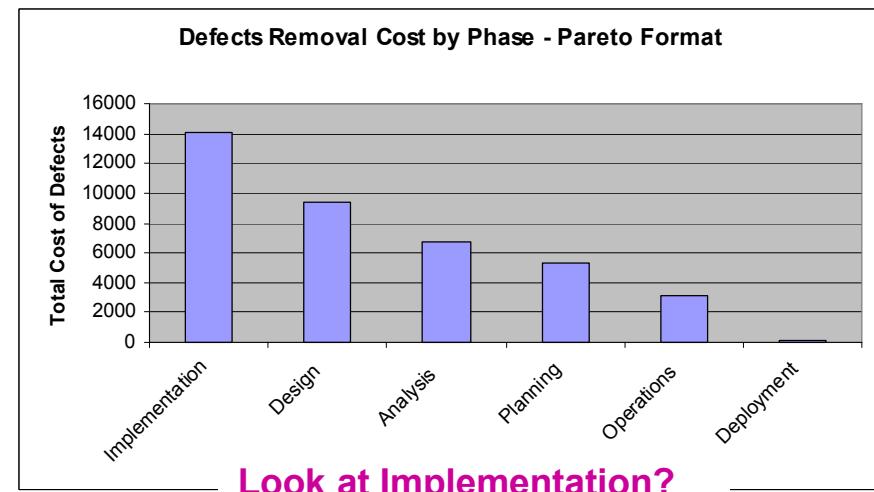
Look at Analysis – first?



Look at Analysis, Design, and Implementation?



Look at Analysis & Design?



Look at Implementation?

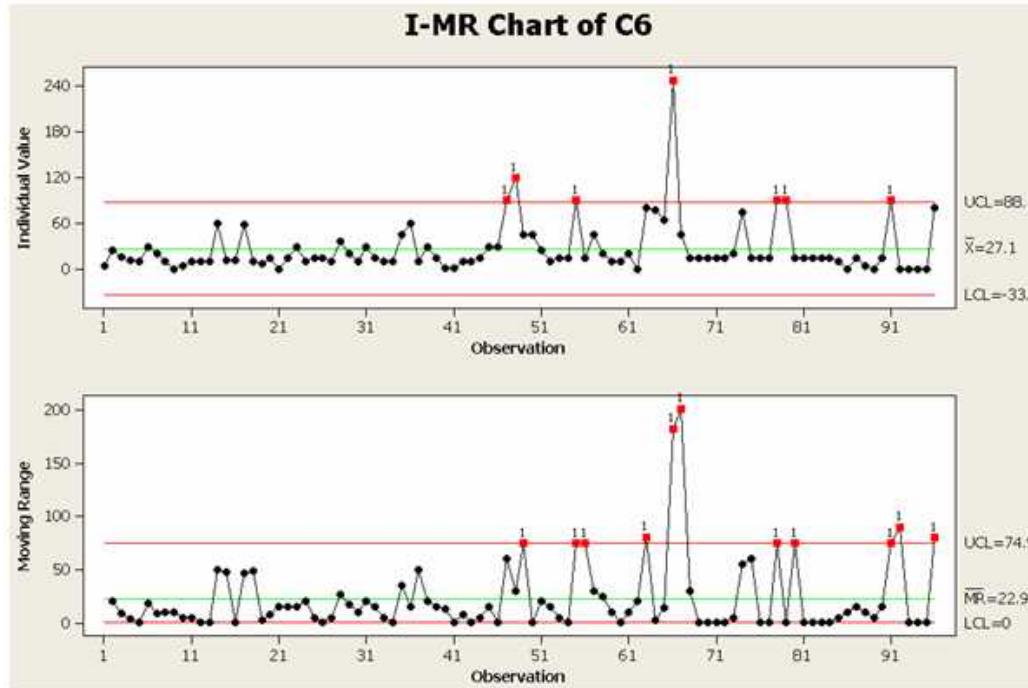
A Way to Look at Defects

		Phase Injected					
		Planning	Analysis	Design	Impl.	Deploy.	Ops.
Phase Detected	Planning	109	4	8	8		
	Analysis	1	290	2			Find
	Design	3	9	476	2		Remove
	Impl.	1	1	13	296		Prevent
	Deploy.				1	20	
	Ops.			3	24	2	30
	Total Injected	114	304	502	331	22	30
% leakage		4	3	3	7	9	

What does this association matrix REVEAL?



Some Statistical Analysis is Required for “Higher Level Maturity”



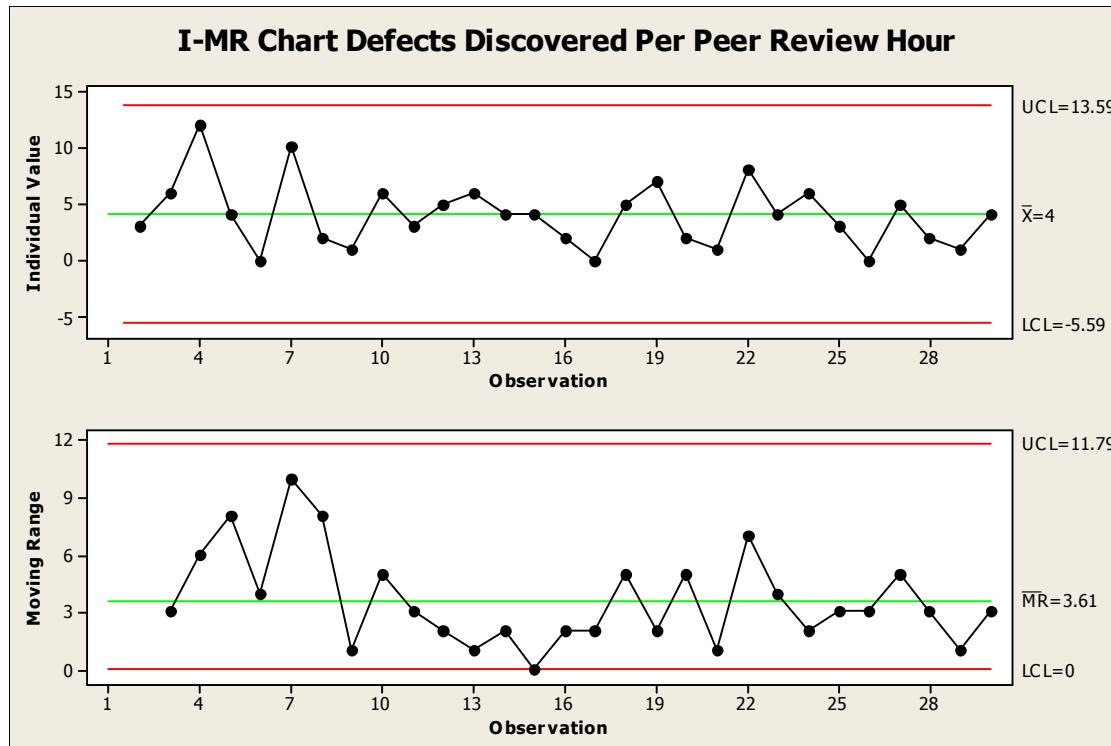
Find
Remove
Prevent

Special (Assignable) Cause removal required at CMMI® Level 4

How well the process is performed

Some Statistical Analysis is Required for “Higher Level Maturity”

Balancing the “voice of the customer” with the “voice of the process”



Causal analysis is required at CMMI® Level 5

Find

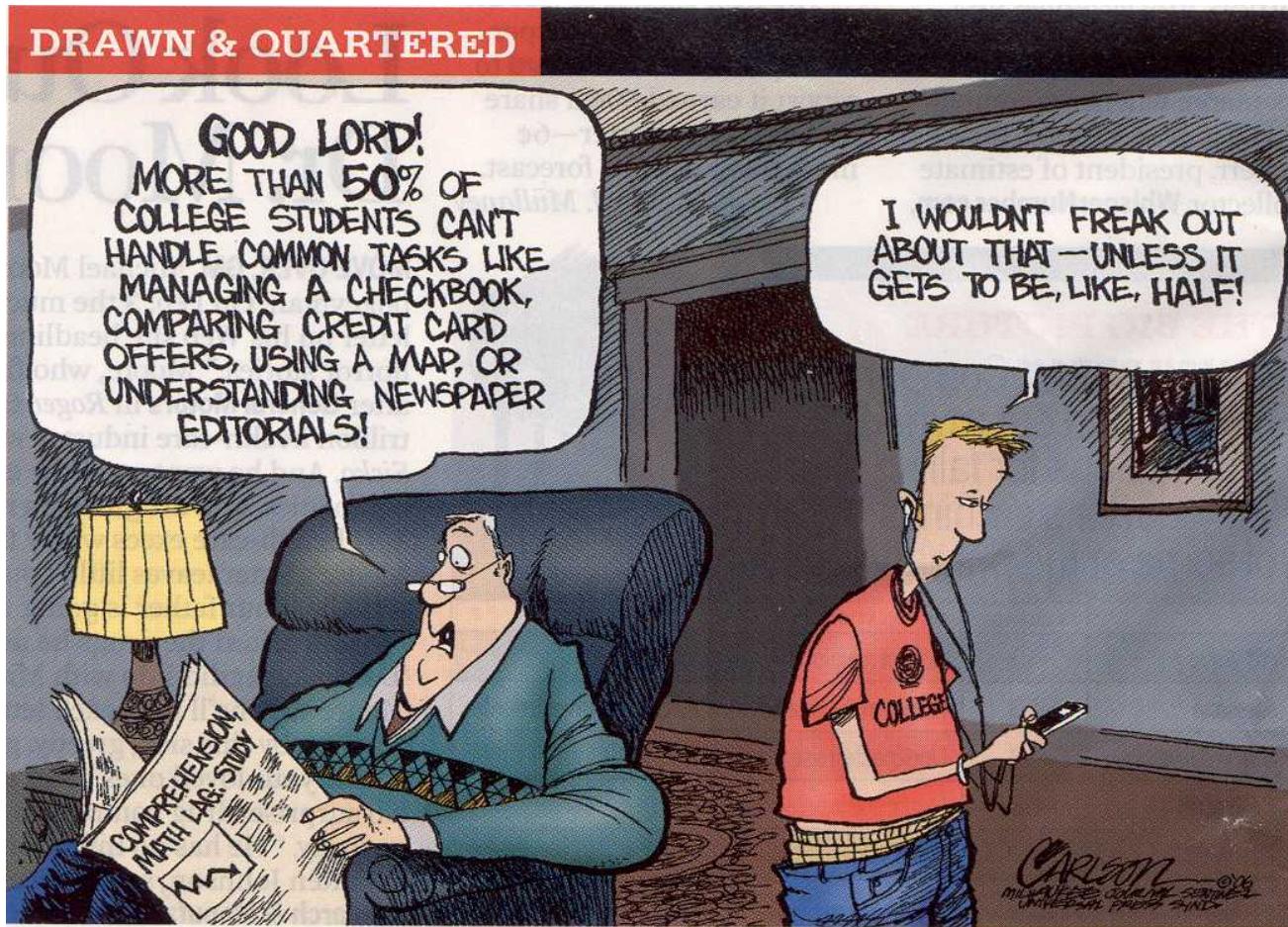
How well the process is performs

Remove

Prevent



Don't worry, the future looks bright – &\$#%!



BusinessWeek 2/20/2006 pg. 14



Measurement-related Characteristics of “Higher Level Maturity” Organizations

- 85 percent of high maturity organization align process improvement with TQM initiatives at the enterprise level
- 85 percent of high maturity organizations have multiple process and quality improvement initiatives
- 81 percent of high maturity organizations also have ISO 9001 certification
- 50 percent using CMM, 42 percent using balanced scorecard, 42 percent using LSS, 25 percent using Baldrige
- 58 percent of high maturity organizations have established formal mentoring programs
- 15 - 21 percent decrease in effort for a one level increase

Higher Level Maturity with Statistics course (SEI)



Measurement-related Characteristics of “Higher Level Maturity” Organizations

Higher level maturity organizations:

- Use centralized measurement
- Know why measurement are used, bring tables and charts to interviews and explain them!
- Describe “causal” analysis spontaneously during assessment interviews
- Have and supply ROI data
- Use Pareto analysis
- Use control charts
- Use Six Sigma
- Use *orthogonal defect classification* – and this means . . .
- Require participation in the SQA Group and / or the SEPG prior to promotion to management

The 2001 High Maturity Workshop (sponsored by the SEI)



Introducing defect pithy 'tudes

(a hostile attitude or disposition; that is, not defect friendly)

Defects persist in software; *most of these come from executing a poor software development process or not executing a good one!*

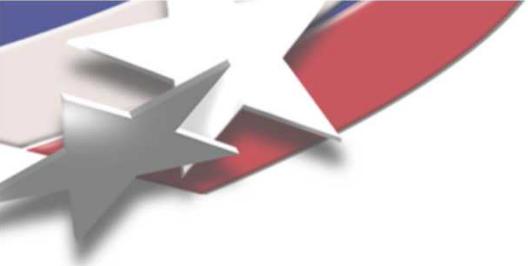
Defectectomy – *Surgical* defect removal, often evidenced in peer reviews

Defecticide – the killing off of defects, often evidenced in testing

Defect metastasization – the absence of Defectectomy and Defecticide practices

$$Q - Dr = F$$

Quality without defect (removal) is merely “faking it.”



Closing Reminders

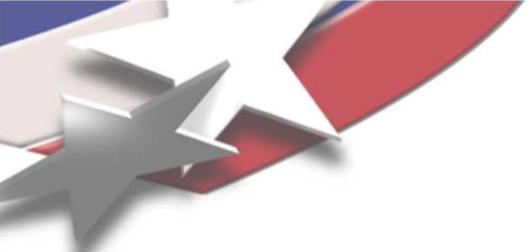
“Higher maturity” organizations MUST use statistical tools for identifying assignable and common cause deviation.

Collecting and statistically analyzing defects is superior to merely counting defects and their origins.

SPC, Pareto, and Histogram charts are simple statistical tools for defect analysis and quantitative improvement.

These same tools can be used in other software development activities (estimated vs. actual, requirements volatility)

The following approaches have been found to be the best in class for defect prevention: JADs during requirements, prototypes during design, and reuse during coding and documentation. Jones, *Software Quality*, pg 160



And we do this because . . .

Defects are an aspect of measurement (and part of the CSMS!)

The customer has already paid for a quality product.

Predicting performance should be based on process capability. (voice of the process)

Most organizations are interested in lower operational costs.

Most defects are preventable.

You won't need to clean-up later and you won't need to be as charming!

